

PROJET DE FIN D'ÉTUDES

Organisme d'accueil : SAHAM ASSURANCE

TARIFICATION ET PROVISIONNEMENT EN ASSURANCE MALADIE GROUPE DE BASE CHEZ SAHAM ASSURANCE

Réalisé par : ENNAJEH Mohammed Amine (Actuariat et finance)
AIT ZEOUAY Soukaina (Actuariat et finance)

Sous la Direction de :

M. CHAOUBI Abdelaziz (INSEA)
Mme ARICHI Zineb
Mlle MAHER Soundous (SAHAM ASSURANCE)

Soutenu publiquement comme exigence partielle en vue de l'obtention du

Diplôme d'Ingénieur d'Etat

Option : ACTUARIAT ET FINANCE

Devant le jury composé de :

- M. CHAOUBI Abdelaziz (Encadrant INSEA)
- M. GUEDIRA Fayçal (Examineur INSEA)
- Mlle. MAHER SOUNDOUS (Encadrant SAHAM ASSURANCE)

REMERCIEMENT

Nous saisissons l'occasion pour témoigner de notre grande reconnaissance envers notre professeur encadrant M. Abdelaziz CHAOUBI que nous remercions pour ses orientations qui nous ont aidées pour rédiger ce projet. Nous espérons que ce projet sera à la hauteur de ses attentes. Nous sommes très reconnaissants à l'égard de Mme Zineb ARICHI (Manager Branche Vie) et Mlle Soundous MAHER (Actuaire), nos deux encadrant au sein de la compagnie SAHAM Assurance, pour leur accueil, leur encadrement au quotidien, leur confiance en nous confiant un projet de telle ampleur, et qui nous a ouvert toutes les voies possibles pour mener à bien notre travail du début jusqu'à la fin. A cette même occasion, nous voudrions remercier aussi le personnel du département Actuariat, qui nous a offert un climat très accueillant. Enfin, nous ne pourrions clore ce passage sans adresser nos remerciements les plus profonds au corps enseignant de l'INSEA, pour le grand effort qu'il fournit afin de nous donner les outils et les connaissances nécessaires pour comprendre les faits, analyser les phénomènes et poser les bonnes questions, et au -delà nous inculquer un savoir -être nécessaire à tout ingénieur.

DEDICACE

Je dédie ce travail,

A mes chers parents

Aucun mot, aucune dédicace ne saurait exprimer mon respect, ma considération et l'amour éternel pour les sacrifices que vous avez consentis pour mon instruction et mon bien être.

Trouvez en ce travail le fruit de votre dévouement et l'expression de ma gratitude et mon profond amour.

A mes frères et sœur,

Vous m'avez toujours soutenu durant toutes mes études, je vous souhaite beaucoup de bonheur et de réussite.

A toute ma famille

A tous mes amies et amis

AIT ZEOUAY Soukaina

Je dédie ce travail aux êtres qui me sont les plus chers :

À mes chers parents qui ne cessent pas de m'encourager et de m'inciter à perpétuer et faire de mon mieux.

À mes adorables frères et sœurs pour leurs affections et leurs prières, ainsi qu'à leurs conjoints,

À Ibtissam Elfakri,

Et à toute ma famille, ainsi qu'à mes amis.

ENNAJEH Mohammed Amine

RESUME

Dans le cadre de l'assurance maladie, compte tenu du progrès de la médecine et face aux développements des maladies graves et à l'augmentation des frais médicaux, les compagnies d'assurance se doivent d'être en mesure de proposer en permanence un tarif adapté au marché, et donc de revoir et mettre à jour de manière fréquente leurs bases de tarification. C'est dans ce contexte que se situe notre projet.

L'objectif de ce mémoire est de proposer une tarification pour chaque groupe assuré à « Saham ». Dans un premier temps, une approche de la tarification de la consommation globale de chaque police par la théorie de la crédibilité a été mise en application, après avoir étudié les caractéristiques et l'historique de la consommation moyenne de chaque groupe. Dans un second temps, une analyse de données a servi pour détecter la corrélation entre les variables explicatives d'une part et pour tester l'égalité de la consommation moyenne par prestation d'autre part. Puis, un modèle du type « fréquence * coût » basé sur des modèles linéaires généralisés a été mis en application pour calculer la prime pure par prestation, qui va nous permettre par la suite d'établir l'outil de tarification exigé par la compagnie. Dans un dernier temps, nous essayerons de calculer les provisions pour sinistres à payer (PSAP) liés à ce portefeuille en utilisant plusieurs méthodes et approches déterministes et stochastiques.

MOTS CLES : Crédibilité, modèle linéaire généralisé, tarification, provisionnement, segmentation, prime pure, prestation

LISTE D'ABREVIATIONS

AMO : Assurance Maladie Obligatoire.

AMC : Assurance Maladie Complémentaire.

CNSS : Caisse National de Sécurité Sociale.

CNOPS : Caisse Nationale des Organismes de Prévoyance Sociale.

RAMED : Régime d'Assistance Médicale.

ANAM : Agence Nationale de l'Assurance Maladie.

TNR: Taux National de Remboursement.

ALD : Affection de Longue Durée.

ALC : Affection de Courte Durée.

GLM : Modèle Linéaire Généralisé.

ANOVA : Analyse de la Variance.

ACP : Analyse en Composante Principale.

ODF : Orthopédie dento-faciale.

LISTE DES TABLEAUX

TABLEAU 1-TYPES DE GARANTIE POUR L'ASSURANCE MALADIE, SOURCE : ANAM.....	14
TABLEAU 2-LA POPULATION IMMATRICULE DE L'AMO (EN PERSONNES), SOURCE :RAPPORT ANNUEL 2013 DE L'AMO	16
TABLEAU 3-PANIER DE SOINS DE L'AMO GERE PAR LA CNOPS, SOURCE : ANAM.....	18
TABLEAU 4-PANIER DE SOINS DE L'AMO GERE PAR LA CNSS, SOURCE : ANAM.....	19
TABLEAU 5-CHIFFRE D'AFFAIRE DE LA MALADIE DES COMPAGNIES D'ASSURANCES (EN MILLE DHS), SOURCE : RAPPORT ANNUEL 2013 DE SAHAM ASSURANCE	20
TABLEAU 6-LA PART DU MARCHE DE LA MALADIE DES DIFFERENTES COMPAGNIES D'ASSURANCE, SOURCE : RAPPORT ANNUEL 2013 DE SAHAM ASSURANCE.....	21
TABLEAU 7: PRESTATIONS OFFERTES PAR SAHAM ASSURANCE	25
TABLEAU 8: CLASSES DE LA VARIABLE EFFECTIF	31
TABLEAU 9: CLASSES DE LA VARIABLE PLAFOND GENERAL	31
TABLEAU 10-LES CLASSES DE TARIFICATION A POSTERIORI	31
TABLEAU 11-LA CONSOMMATION MOYENNE PAR CLASSE SUR LES QUATRE ANNEES D'ETUDES (EN DHS).....	32
TABLEAU 12- LES PARAMETRES DE STRUCTURE DU MODELE DE BÜHLMANN.....	33
TABLEAU 13-EXEMPLES DE PRIME DE CREDIBILITE POUR LA 5eme CLASSE	33
TABLEAU 14-LES PARAMETRES DE STRUCTURE DU MODELE BÜHLMANN-STRAUB.....	34
TABLEAU 15-STATISTIQUE DE LEVENE	36
TABLEAU 16-STATISTIQUE DE WELCH	37
TABLEAU 17- RESULTAT DU TEST D'EGALITE DE LA CONSOMMATION MOYENNE PAR RAPPORT A LA PRESTATION « ANALYSE (BIOLOGIE) »	38
TABLEAU 18: L'INERTIE EXPLIQUEE PAR CHAQUE AXE FACTORIELLE.....	41
TABLEAU 19: LE SCORE DE CHAQUE VARIABLE SUR LES AXES FACTORIELS RETENUS.....	42
TABLEAU 20: CODIFICATION DE NOS VARIABLES TARIFAIRES.....	46
TABLEAU 21-REPARTITION DES EFFECTIFS PAR TYPE DE BENEFICIAIRE.....	46
TABLEAU 22- AGE MOYEN PAR TYPE DE BENEFICIAIRE.....	47
TABLEAU 23-REPARTITION DES EFFECTIFS PAR CLASSE DU TAUX DE REMBOURSEMENT	47
TABLEAU 24-REPARTITION DES EFFECTIFS PAR CLASSE DU PLAFOND GENERAL.....	48
TABLEAU 25-REPARTITION DES FRAIS ENGAGES PAR TYPE DE BENEFICIAIRE	48
TABLEAU 26-REPARTITION DES FRAIS ENGAGES PAR PRESTATION	49
TABLEAU 27- COUT MOYEN ET FREQUENCE PAR TYPE DE BENEFICIAIRE	57
TABLEAU 28-COUT MOYEN ET FREQUENCE PAR TAUX DE REMBOURSEMENT	58
TABLEAU 29-COUT MOYEN ET FREQUENCE PAR PLAFOND GENERAL.....	58

TABLEAU 30-COUT MOYEN ET FREQUENCE PAR EFFECTIF DE GROUPE	59
TABLEAU 31- STATISTIQUE DE DEVIANCE DU MODELE PAR LOI	60
TABLEAU 32-STATISTIQUES TYPE 3 DES VARIABLES DU MODELE AJUSTE AU COUT PAR UNE LOI LOGNORMALE	61
TABLEAU 33-LES COEFFICIENTS ESTIMES DU MODELE DU COUT DES ASSURES POUR LA « PHARMACIE »	62
TABLEAU 34-TEST DE CONTRASTE POUR LE REGROUPEMENT DES CLASSES DES MODALITES DE LA VARIABLE EFFECTIF	62
TABLEAU 35- STATISTIQUES TYPE 3 DES VARIABLES DU MODELE AJUSTE AU COUT PAR UNE LOI LOGNORMALE APRES REGROUPEMENT DES MODALITES.....	62
TABLEAU 36-LES NOUVEAUX PARAMETRES ESTIMES APRES REGROUPEMENT DES MODALITES	63
TABLEAU 37-STATISTIQUE DE DEVIANCE DU MODELE DE FREQUENCE PAR LOI	64
TABLEAU 38- STATISTIQUES TYPE 3 DES VARIABLES DU MODELE AJUSTE DE LA FREQUENCE PAR UNE LOI BINOMIALE NEGATIVE.....	64
TABLEAU 39-PARAMETRES ESTIMES DU MODELE DE FREQUENCE POUR « LA PHARMACIE ».....	65
TABLEAU 40-TEST SUR LE REGROUPEMENT DES MODALITES DE LA VARIABLE TAUX DE REMBOURSEMENT	65
TABLEAU 41-TEST SUR LE REGROUPEMENT DES MODALITES DE LA VARIABLE PLAFOND GENERAL	65
TABLEAU 42-STATISTIQUES DE TYPE 3 APRES REGROUPEMENT DES MODALITES	66
TABLEAU 43-LES NOUVEAUX PARAMETRES ESTIMES APRES REGROUPEMENT DES MODALITES	66
TABLEAU 44-CARACTERISTIQUES DE LA CLASSE DE REFERENCE	67
TABLEAU 45-LES VALEURS PROPRES DES AXES FACTORIELS.....	74
TABLEAU 46-LES INDICATEURS DE QUALITE DE CLASSIFICATION "K-MEANS"	75
TABLEAU 47-REPARTITION DE LA CONSOMMATION PAR PRESTATION (EN % DE LA CONSOMMATION TOTALE DE CHAQUE CLASSE)	77
TABLEAU 48-CARACTERISTIQUES SOCIODEMOGRAPHIQUES DES CLASSES D'INDIVIDUS.....	78
TABLEAU 49-TRIANGLE DES REGLEMENTS DES PAIEMENTS.....	81
TABLEAU 50-FACTEURS DE DEVELOPPEMENT DE LA METHODE CHAIN LADDER	84
TABLEAU 51 : RESERVES ET ECARTS TYPES ESTIMES PAR LA METHODE CHAIN LADDER	85
TABLEAU 52: RESERVES ET ECARTS TYPES ESTIMES PAR LA METHODE MACK.....	85
TABLEAU 53: COEFFICIENTS ESTIMES PAR LA METHODE DE VYLDER	85
TABLEAU 54-CRITERE DE VALIDATION DU MODELE.....	85
TABLEAU 55: RESERVES ET ECARTS TYPES ESTIMES PAR LA METHODE DEVYLDER (EN DHS)	86
TABLEAU 56: RESERVES ET ECARTS TYPES ESTIMES PAR LA METHODE BOOTSTRAP (EN DHS).....	86
TABLEAU 57: RESERVES ET ECARTS TYPES ESTIMES PAR LA METHODE GLM (EN DHS)	86
TABLEAU 58- COMPARAISON DES METHODES POUR LES RESERVES ESTIMEES (EN DHS)	86
TABLEAU 59-DIFFERENCE D'ESTIMATION DES RESERVES (EN DHS)	86
TABLEAU 60: COMPARAISON DES RESERVES DES DIFFERENTES METHODES (EN DHS)	87
TABLEAU 61 : ESTIMATION ET INTERVALLE DE CONFIANCE DES RESERVES PAR LA METHODE MACK (EN DHS)	87

LISTE DES FIGURES

FIGURE 1- EVOLUTION DE L'AMO.....	15
FIGURE 2- POPULATION DES COTISANTS A L'AMO PAR SECTEUR (EN MILLE PERSONNES), SOURCE : RAPPORT ANNUEL 2013 DE L'AMO	16
FIGURE 3: POPULATION DES AYANTS DROITES A L'AMO PAR SECTEUR (EN MILLE PERSONNES), SOURCE : RAPPORT ANNUEL 2013 DE L'AMO	17
FIGURE 4- EVOLUTION DU CHIFFRE D'AFFAIRE DE L'ASSURANCE MALADIE ET CELUI DE SAHAM ASSURANCE (EN MILLE DHS), SOURCE : RAPPORT ANNUEL 2013 DE SAHAM ASSURANCE	21
FIGURE 5-L'EVOLUTION DE PART DE MARCHE SAHAM ASSURANCE.....	22
FIGURE 6-DIAGRAMME DES COMPOSANTES DANS L'ESPACE.....	42
FIGURE 7-ARBRE DE CODIFICATION DU PLAFOND GENERAL	45
FIGURE 8-ARBRE DE CODIFICATION NAISSANCE SIMPLE.....	45
FIGURE 9-REPARTITION DES EFFECTIFS PAR TYPE DE BENEFICIAIRES.....	47
FIGURE 10-QQ-PLOT DES LOIS GAMMA ET LOG NORMALE	60
FIGURE 11-REPRESENTATION GRAPHIQUE DES RESIDUS DE LA DEVIANCE.....	63
FIGURE 12: BOUTON DEMARRAGE D'APPLICATION	68
FIGURE 13: FENETRE "SAISIE DES DONNEES"	69
FIGURE 14: FENETRE "TARIFICATION"	70
FIGURE 15- LE DENDROGRAMME DES AGREGATIONS SUCCESSIVES	75
FIGURE 16-GRAPHE DU R CARRE SEMI PARTIEL EN FONCTION DU NOMBRE DE CLASSES	76

TABLE DES MATIERES

CHAPITRE 1 : PRESENTATION DE L'ASSURANCE MALADIE AU MAROC	13
I. L'ASSURANCE MALADIE AU MAROC :.....	13
I.1) DEFINITION DE QUELQUES CONCEPTS :.....	13
I.2) L'ASSURANCE MALADIE OBLIGATOIRE AU MAROC :.....	14
II. SAHAM ASSURANCE DANS LE MARCHÉ DE LA COUVERTURE MALADIE :	20
II.1) EVOLUTION DU CHIFFRE D'AFFAIRE DE DIFFÉRENTES COMPAGNIES D'ASSURANCE :.....	20
II.2) EVOLUTION PART DE MARCHÉ ASSURANCE MALADIE :.....	21
CHAPITRE 2 : PRESENTATION DES DONNEES DU PORTEFEUILLE.....	23
I. LA STRUCTURE DES BASES UTILISEES :	23
II. LA CONSTITUTION DES BASES DE DONNEES :	24
III. REGROUPEMENT DES VARIABLES :.....	24
IV. CREATION DES VARIABLES :.....	25
CHAPITRE 3 : TARIFICATION A POSTERIORI	26
I. LA THEORIE DE CREDIBILITE :.....	26
I.1) PRESENTATION DU MODELE DE BÜHLMANN :.....	27
I.2) PRESENTATION DU MODELE DE BÜHLMANN –STRAUB :	29
I.3) APPLICATION DE LA THEORIE DE CREDIBILITE SUR LA POPULATION ETUDIEE :	30
II. ANALYSE DE LA VARIANCE :.....	35
II.1) TEST D'EGALITE DES VARIANCES :.....	35
II.2) TEST D'EGALITE DES MOYENNES :.....	36
CHAPITRE 4 : TARIFICATION A PRIORI	40
I. TAUX DE REMBOURSEMENT PONDERE :	40
II. L'ANALYSE EN COMPOSANTES PRINCIPALES :	40
III. SEGMENTATION ET CODIFICATION DES VARIABLES :.....	43
III.1) PRESENTATION DE L'ALGORITHME CHAID :.....	43
III.2) SEGMENTATION DES VARIABLES :.....	44

IV.	STATISTIQUES DESCRIPTIVES SUR LE PORTEFEUILLE :	46
V.	LE CADRE THEORIQUE DU MODELE LINEAIRE GENERALISE :	50
V.1)	RAPPEL DU MODELE LINEAIRE SIMPLE :	50
V.2)	DISTRIBUTION DE LA FAMILLE EXPONENTIELLE NATURELLE :	50
V.3)	HYPOTHESE DU MODELE LINEAIRE GENERALISE :	53
V.4)	ESTIMATION DES PARAMETRES :	54
V.5)	SIGNIFICATIVITE DES VARIABLES :	54
V.6)	QUALITE DU MODELE :	55
V.7)	SELECTION DES VARIABLES :	55
V.8)	VALIDATION DU MODELE :	56
VI.	APPLICATION DE LA TARIFICATION A L'AIDE D'UN MODELE LINEAIRE GENERALISE :	56
VI.1.	MODELISATION DE LA GARANTIE « PHARMACIE » :	57
VI.2.	METHODE DU CALCUL DU TARIF GLOBAL :	67
VII.	OUTIL DE TARIFICATION :	68
CHAPITRE 5 : ETUDE COMPORTEMENTALE DE LA CONSOMMATION MEDICALE		71
I.	PRESENTATION DES METHODES DE CLASSIFICATION :	71
I.1)	LA METHODE DE PARTITIONNEMENT « K-MEANS » :	71
I.2)	LA METHODE DE CLASSIFICATION ASCENDANTE HIERARCHIQUE :	71
I.3)	LA CLASSIFICATION MIXTE :	72
I.4)	CHOIX DU NOMBRE DE GROUPES DE LA METHODE CAH :	72
II.	CLASSIFICATION DES ASSURES DU PORTEFEUILLE ETUDIEE :	73
II.1)	REDUCTION ET HOMOGENEISATION DES DONNEES :	74
II.2)	APPLICATION DE LA METHODE DE PARTITIONNEMENT « K-MEANS » :	74
II.3)	APPLICATION DE LA METHODE DE CLASSIFICATION ASCENDANTE HIERARCHIQUE (CAH) :	75
CHAPITRE 6 : PROVISIONNEMENT		80
I.	LA VIE DYNAMIQUE DE VIE D'UN SINISTRE :	80
II.	LA METHODE DE CHAIN LADDER :	81
III.	MODELE DE MACK :	82
IV.	MODELE DE VYLDER ET DE CHRISTOPHIDES :	83
V.	BOOTSTRAP:	83

VI.	GLM :	84
VII.	APPLICATION DES METHODES DE PROVISIONNEMENT :	84
	CONCLUSION GENERALE	88
	BIBLIOGRAPHIE	89
	ANNEXE 1:	ERREUR ! SIGNET NON DEFINI.
	ANNEXE 2:	ERREUR ! SIGNET NON DEFINI.
	ANNEXE 3:	ERREUR ! SIGNET NON DEFINI.
	ANNEXE4:	

INTRODUCTION

L'assurance Maladie, connaît depuis plusieurs années une succession de réformes ayant pour objectif principal d'apporter des réponses aux difficultés financières de la Sécurité Sociale.

Dans ce contexte, l'état a fait appel aux compagnies d'assurance afin de jouir de leurs expériences en termes d'appréciation et de maîtrise du risque Santé. Ainsi ces compagnies d'assurance ont pris la relève dans le domaine de la couverture maladie.

Une bonne gestion du risque santé par les assureurs passe nécessairement par une actualisation fréquente de leurs tarifs appliqués afin de mieux se couvrir contre les risques assurés.

Suite aux phénomènes liés à la branche Maladie à savoir le risque croissant et l'aléa moral, le portefeuille maladie de SAHAM Assurance devient déficitaire à cause des tarifs appliqués actuellement et qui sont dangereusement arrêtés par des méthodes statistiques non appropriées.

Dans le but de rééquilibrer leurs portefeuilles Maladie, SAHAM Assurance a lancé un projet de révision de leurs tarifs appliqués, c'est à cet effet qu'ils ont fait appel à nous dans le but d'arrêter de nouveaux tarifs calculés par des méthodes statistiques.

Le présent mémoire expose la méthodologie adoptée afin d'aboutir à une réalisation du dit projet. Nous avons fait ainsi appel à nos connaissances actuarielles pour bien mener ce travail.

Notre étude ne portera que sur la tarification et le provisionnement de l'assurance groupe de base, c'est-à-dire remboursement du 'premier Dirhams engagé'. Ainsi nous avons filtré notre base de données afin d'isoler la population ciblée par cette refonte de tarification à savoir les personnes actives et ayants souscrits à une couverture de groupe.

Premièrement, nous avons calculé la consommation totale du groupe afin d'appliquer la théorie de crédibilité et arrêter un tarif portant sur la consommation totale. Une analyse de l'égalité de la moyenne a été ajoutée à cette étape afin de vérifier l'hypothèse de différence de moyenne de consommation entre les différents types de prestations offerts par la compagnie.

Cette étude d'égalité des moyennes nous a permis de dégager une différence significative entre la consommation des différentes prestations, ainsi une tarification plus fine et plus segmentée a été réalisée en s'appuyant sur le modèle linéaire généralisé. De prime abord nous avons réalisé une analyse en composante principale et ce dans le but d'étudier et analyser nos différentes variables tarifaires. Ensuite la tarification effectuée a permis de concevoir un outil de tarification apte à être implémenté chez les différents courtiers et agences offrants les services de SAHAM Assurance.

Enfin, Nous avons estimés par différentes méthodes de provisionnement les réserves, et spécialement les 'Provisions Pour Sinistre à Payer' afin de permettre à la compagnie dans un premier temps de concevoir une idée sur les règlements futurs et dans un second temps d'étudier les Ratio de sinistralité $\left(\frac{S}{P}\right)^1$.

¹ $\left(\frac{S}{P}\right) = \frac{\text{Total des Sinistres}}{\text{Total des Primes}}$.

CHAPITRE 1 : PRESENTATION DE L'ASSURANCE MALADIE AU MAROC

Dans ce chapitre, nous allons donner un aperçu sur le marché d'assurance maladie au Maroc. Tout d'abord nous allons présenter le produit qu'offrent les assurances privées pour la maladie, puis nous allons détailler la couverture médicale de base et finalement comparer l'évolution du chiffre d'affaire de la maladie de SAHAM Assurance Maroc par rapport à celle du marché des assurance.

I. L'Assurance Maladie au Maroc :

L'assurance maladie est une assurance contre les dépenses de soin de santé qui a pour objet de couvrir l'assuré en cas de survenance d'une maladie, d'un accident ou d'un accouchement entraînant des frais d'hospitalisation comme les nuitées, les honoraires médecins, produits pharmaceutiques, etc.

I.1) Définition de quelques concepts :

1) La garantie :

L'assurance maladie couvre les frais médicaux, pharmaceutiques et d'hospitalisation ou une partie de ces risques. Outre le fait que la maladie ne doit pas être due à un fait intentionnel, l'assuré ou l'assureur peuvent exclure certains risques de la garantie, par exemple : maladie suite à l'usage de stupéfiants, alcools, compétitions sportives etc. En plus des exclusions, le contrat peut prévoir des franchises. L'assurance contre la maladie est souscrite soit à titre principal soit en complément d'une assurance contre les accidents du travail notamment. L'assuré est tenu de faire une déclaration précise du risque et, comme pour l'assurance vie, il lui est remis un questionnaire sur son état de santé.

2) Le sinistre :

La déclaration d'un sinistre est soumise aux règles générales applicables pour les assurances de dommage. L'assuré est tenu de faire connaître à son assureur le nom de son médecin traitant, le contrat peut prévoir aussi une contre visite médicale. L'assuré est également tenu de communiquer à son assureur les factures et notes d'honoraires. L'assurance des frais médicaux, pharmaceutiques et d'hospitalisation, étant une assurance dommage, elle est donc soumise aux règles applicables en matière de subrogation (assureur possède un droit de recours contre le responsable du dommage et de résiliation pour sinistre). L'assureur pourra résilier le contrat par suite de sinistre auquel cas l'assuré est en droit de résilier ces autres risques garantis par l'assureur.

Le tableau 1 contient les types de garantie pour l'assurance maladie :

Types de garantie pour l'assurance maladie						
Consultation et visite	Médicament	Laboratoire et radiologie	Hospitalisation	Optique	Soin prothèse et parodontie	Maternité

Toutes consultations de nuit ou de dimanche.	Tous frais pharmaceutiques pour les médicaments.	Toutes analyses médicales et de Radiologie.	Toutes hospitalisations chirurgicales dans les cliniques privées, dans les hôpitaux publics ou les polycliniques et toutes hospitalisations médicales comme les honoraires médecins et frais de séjour.	Tous frais de verres et de montures.	Tous soins conservateurs et les soins chirurgicaux.	Tous frais d'accouchement normale ou suite à une césarienne .
--	--	---	---	--------------------------------------	---	---

Tableau 1-Types de garantie pour l'assurance maladie, source : ANAM

I.2) L'assurance maladie obligatoire au Maroc :

Le régime marocain d'assurance maladie couvre les salariés du secteur public et ceux du secteur privé. Il assure aux intéressés une protection contre les risques de différentes maladies.

On distingue trois types de contrats d'assurance maladie qui sont :

- **Contrat assurance maladie obligatoire(AMO) :** Ce régime est entré en vigueur le 18 août 2005 a été institué en 2002 par la loi 65.00² portant code de la couverture médicale, cette assurance garantit les prestations qui comprennent en général les types de frais qu'on a cité dans la maladie.
- **Contrat assurance maladie complémentaire(AMC) :** Elle garantit une couverture des soins exclus de l'AMO comme elle présente un complément du taux de remboursement des soins de l'AMO.
- **Contrat assurance de base :** C'est un contrat maladie de base proposé pour les entreprises privées ayant souscrit un contrat maladie -au niveau des compagnies d'assurance- avant la mise en place de l'AMO.

Ainsi, Le régime d'assurance maladie obligatoire est géré par les organismes suivants :

- **La Caisse Nationale de Sécurité Sociale (CNSS) :** elle gère l'assurance maladie obligatoire du secteur privé.
- **La caisse nationale des organismes de prévoyance sociale (CNOPS) :** elle gère l'assurance maladie obligatoire du secteur public.
- **Les compagnies d'assurance :** elles gèrent de l'assurance maladie obligatoire (article 114³ de la loi 65.00) ainsi que l'assurance maladie complémentaire du secteur public et privé.

Le régime d'assistance médicale (RAMED) est géré par les organismes suivants :

- **l'état et les collectivités locales :** ils couvrent les personnes démunies non couvertes par le régime d'assurance maladie obligatoire de base.

² Veuillez voir l'Annexe2

³ Veuillez vois l'Annexe2

Pour veiller au bon fonctionnement de la couverture médicale de base de ces organismes, l'agence nationale de l'assurance maladie (ANAM) est faite pour cela et a pour mission le contrôle et l'encadrement technique de l'AMO.

1) Généralités sur l'Assurance Maladie Obligatoire :

✓ Evolution de l'AMO :

L'AMO a connue plusieurs évolutions depuis sa mise en place, la figure 1 retrace ces évolutions.

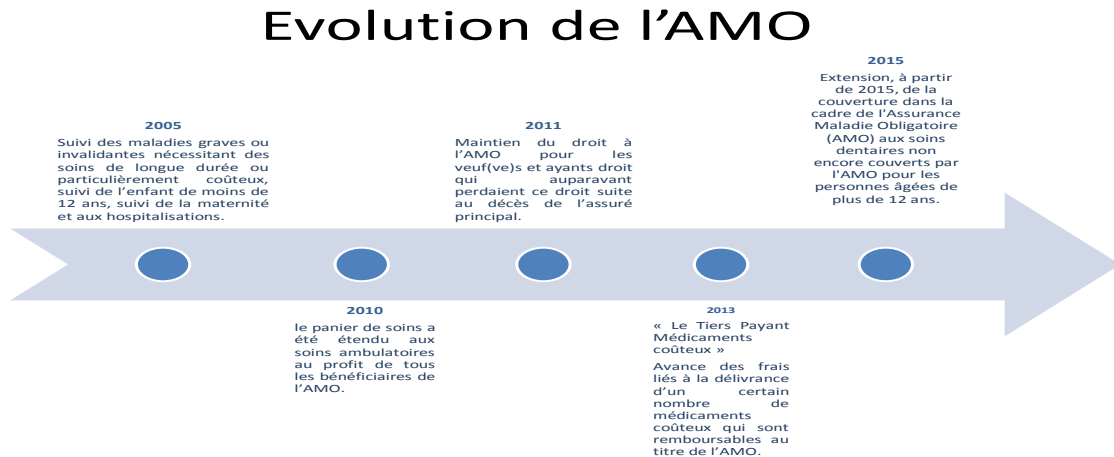


Figure 1- Evolution de l'AMO

✓ Chiffres clés concernant l'AMO

Ces Statistiques ont été tirées du Rapport annuel globale de l'AMO 2013.

1) Population de l'AMO :

- 8 Millions de bénéficiaires de l'AMO.
- 5,1 Millions de bénéficiaires au secteur privé.
- 2,9 Millions de bénéficiaires au secteur public.
- Evolution annuelle moyenne de 6,8% entre 2006 et 2013.
- Age moyen des bénéficiaires du privé et du public est respectivement de 28.7 ans et de 34,8 ans.

2) Dépenses de l'AMO :

- 5,5 Milliards de Dhs de dépenses en 2013.
- 32,3% des dépenses sont destinées aux médicaments.
- 49,3% est la part des dépenses ALD.
- Evolution annuelle moyenne de 15,3% entre 2006 et 2013.

3) Ressources de l'AMO :

- 8,6 Milliards de Dhs de cotisations et contributions en 2013.
- 7,8 Milliards de cotisations et contributions des salariés actifs en 2013.
- 774 Millions de cotisations des pensionnés en 2013.
- Evolution annuelle moyenne de cotisations et contributions de 9,1% entre 2006 et 2013.

✓ Evolution de la population immatriculée de l'AMO :

Le tableau 2 détaille l'évolution de la population immatriculée de l'AMO par type de bénéficiaire et par état d'activité entre 2009 et 2014 :

	Année				
	2009	2010	2011	2012	2013
Actifs	5 123 058	5 327 559	5 479 495	6 209 250	6 555 233
Assurés	2 057 485	2 182 123	2 346 105	2 703 816	2 819 625
Ayants droits	3 065 573	3 145 436	3 133 390	3 505 434	3 735 608
Pensionnées	1 247 711	1 316 068	1 277 679	1 423 169	1 490 396
Assurés	611 744	644 264	675 775	727 034	770 124
Ayants droits	635 967	671 804	601 904	696 135	720 272
Total	6 370 769	6 643 627	6 757 174	7 632 419	8 045 629

Tableau 2-La population immatriculé de l'AMO (en personnes), Source :Rapport Annuel 2013 de l'AMO

D'après le tableau précédent, nous constatons que la population immatriculée de l'AMO ne cesse de croître pour atteindre un total de 8 045 629 personnes. La progression est beaucoup plus importante pour la population du secteur privé qui est passée de 3 741 887 à 5 063 325 personnes entre 2009 et 2013 contrairement à la population du secteur public qui est passée à peine de 2 628 882 en 2009 à 2 982 304 en 2013.

La figure 2 ventile la population des cotisants à l'AMO de base selon le secteur d'activité :

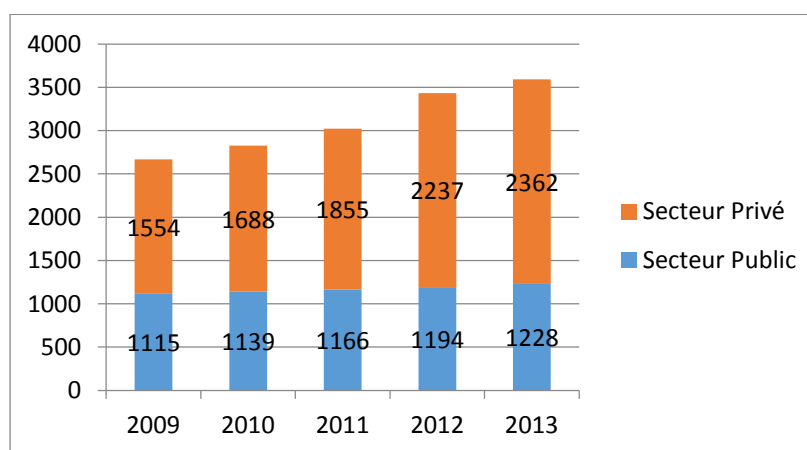
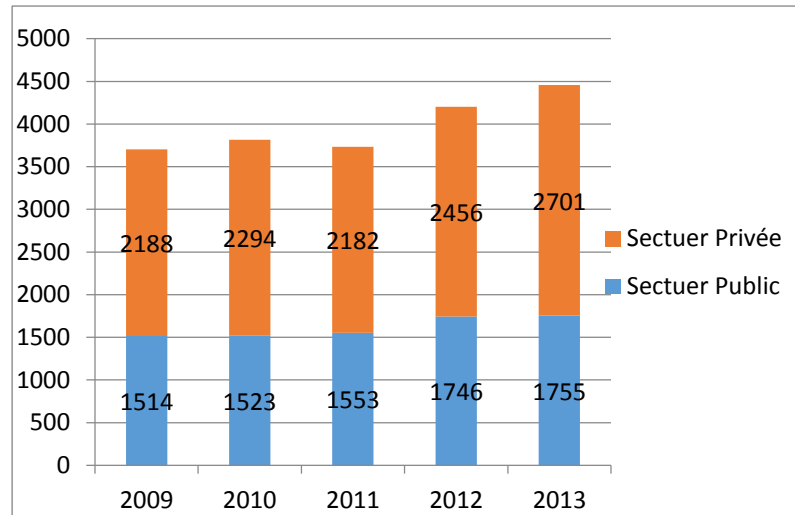


Figure 2- Population des cotisants à l'AMO par secteur (en mille personnes), Source : Rapport Annuel 2013 de l'AMO

Le nombre des cotisants à l'AMO de base est passé de 2 669 229 en 2009 à 3 589 749 personnes en 2013. Au sein de cette population, le poids des salariés du secteur privé représente 66% en 2013 contre 58% en 2009.

La figure 3 quant à elle ventile la population des ayants droits à l'AMO de base selon le secteur :



**Figure 3: Population des ayants droits à l'AMO par secteur (en mille personnes),
Source : Rapport Annuel 2013 de l'AMO**

En termes de population des ayants droit, le poids des bénéficiaires de la couverture dans le secteur privé est moins important, 61% en 2013 contre un poids des assurés de 66%, alors que dans le secteur public, le poids des assurés est de 34% et celui de la population des ayants droit est de 39%: les assurés du secteur privé génèrent moins d'ayants droit que les assurés du secteur public.

2) Assurance Maladie Obligatoire géré par les CNOPS :

✓ La population éligible :

Les agents du secteur public sont soumis obligatoirement à l'assurance maladie obligatoire, ils sont :

- Les actifs : les fonctionnaires et les agents de l'Etat, des collectivités locales, des établissements publics et des personnes morales de droit public.
- Les retraités : les titulaires de pensions du secteur public (y compris les pensions de reversions).
- Les ayant droits : les handicapés à vie, les enfants à l'âge de 21 ans, les étudiants jusqu'à l'âge de 26 ans et les conjoints.

✓ Le Panier de soins :

Dans le panier de soins, nous avons cité les types de soins que CNOPS couvre ainsi que les prestations qu'elle donne, ceci est présenté dans le tableau 3 suivant :

<i>Le panier de soins</i>	
<i>Les types de soins</i>	<i>Les prestations</i>
<p>- Actes de médecine générale et de spécialités médicales et chirurgicales, actes paramédicaux, de rééducation fonctionnelle et de kinésithérapie délivrés à titre ambulatoire hors médicaments.</p> <p>- Soins liés à l'hospitalisation et aux interventions chirurgicales y compris les actes de chirurgie réparatrice et le sang et ses dérivés labiles.</p> <p>- Médicaments admis au remboursement.</p> <p>- Lunetterie médicale, dispositifs médicaux et implants nécessaires aux actes médicaux et chirurgicaux.</p> <p>- Appareils de prothèse et d'orthèse médicales admis au remboursement.</p> <p>- Soins bucco-dentaires.</p> <p>- Orthodontie médicalement requise pour les enfants.</p>	<p>ALD et ALC : 100 % sur la base de la tarification nationale de référence (TNR).</p> <p>- Hospitalisation dans les hôpitaux publics : 100 % de la TNR.</p> <p>- Hospitalisation dans les cliniques privées : 90 % de la TNR.</p> <p>- Médicaments : 70 % (Le remboursement des dépenses de médicaments s'effectue sur la base du PPM du médicament générique, lorsqu'il existe.)</p> <p>- Soins ambulatoires : 80 % de la TNR.</p> <p>- Appareillage, lunetterie, prothèses, orthèses, implants et dispositifs médicaux : sous forme de forfaits prévus dans la TNR.</p>

Tableau 3-Panier de soins de l'AMO géré par la CNOPS, Source : ANAM

✓ Le taux de cotisation

Le taux de cotisation pour l'assurance maladie n'est pas le même pour les actifs que pour les pensionnés du secteur public, on a donc pour :

- Les actifs : Le taux de cotisation est de 5% de l'ensemble des rémunérations, 2,5 % à la charge de l'employeur et 2,5 % qui reste est à la charge du salarié.
- Les retraités : Le taux de cotisation est de 2,5% du montant global des pensions de base.

Remarque : Pour la CNOPS, toutes les cotisations doivent atteindre un seuil minimal mensuel de 70 DH et plafonné à 400 DH.

3) Assurance Maladie Obligatoire géré par la CNSS :

✓ La population éligible

Les agents du secteur privé sont soumis obligatoirement à l'assurance maladie obligatoire de la CNSS, ils sont :

- Les actifs : les salariés du secteur privé.
- Les retraités : les titulaires de pensions du secteur privé.
- Les ayants droits, les handicapés à vie, l'enfant à l'âge de 21 ans, les étudiants jusqu'à l'âge de 26 ans et les conjoints.

✓ Le panier de soins

Dans le panier de soins, nous avons cité les types de soins que la CNSS couvre ainsi que les prestations qu'elle donne à ses assurés et ceci est présenté dans le tableau 4 :

<i>Le panier de soins</i>	
<i>Les types de soins</i>	<i>Les prestations</i>
-ALD et ALC - Suivi de la grossesse, l'accouchement et ses suites, les actes médicaux et chirurgicaux. - Les médicaments admis au remboursement, le sang et ses dérivés labiles, les actes paramédicaux et les actes de rééducation fonctionnelle et de kinésithérapie. - Hospitalisation : l'ensemble des prestations et soins rendus dans ce cadre y compris les actes de chirurgie réparatrice. - Soins Ambulatoires.	-90% pour les prestations assurées par les professionnels et établissements de soins du secteur public (TNR). - 70% pour les prestations assurées par les professionnels et établissements de soins du secteur privé (TNR). - 90% pour les ALD et ALC (TNR).

Tableau 4-Panier de soins de l'AMO géré par la CNSS, Source : ANAM

✓ Le taux de cotisation

Le taux de cotisation pour l'AMO change selon les actifs et les pensionnés du secteur privé, on a donc pour :

- Les actifs : le taux de cotisation est de 4% de l'ensemble des rémunérations réparties en deux, 2 % à la charge de l'employeur et 2 % à la charge du salarié qui est majoré de 1,50% de l'ensemble de la rémunération brute mensuelle du salarié.
- Les retraités : le taux de cotisation due par les titulaires de pensions est fixé à 4% sur le montant global des pensions de base.

Remarque : Contrairement à la CNOPS, les cotisations pour la maladie à la CNSS ne sont pas plafonnées.

4) Assurance Maladie Obligatoire gérée par les compagnies d'assurances :

L'assurance maladie obligatoire qui est gérée par les assurances concerne les salariés du secteur privé, les fonctionnaires de l'état et les titulaires de pensions du secteur privé et public qui sont connus par l'article 114 ainsi que leurs ayants droits qui sont les handicapés à vie, les enfants à l'âge de 21 ans, les étudiants jusqu'à l'âge de 26 ans et leurs conjointes. Les compagnies d'assurance couvrent tous les types de soins qui sont couverts par la CNSS et la

CNOPS avec des prestations qui varient généralement de 70% à 90% des charges selon le panier choisi par l'assuré en contrepartie d'une cotisation qui est soit forfaitaire soit avec un taux de cotisation sur l'ensemble de la masse salariale.

II. SAHAM Assurance dans le marché de la couverture maladie :

La compagnie d'assurance Saham Assurance est présentée à l'annexe 1.

II.1) Evolution du chiffre d'affaire de différentes compagnies d'assurance :

Le tableau 5 contient les données annuelles s'étalant de 2008 à 2013 du chiffre d'affaire de la branche maladie des différentes compagnies d'assurances. On constate que SAHAM Assurance se positionne en pôle position en termes de chiffre d'affaire en 2013.

	2008	2009	2010	2011	2012	2013
Saham Assurance	125 041	450 065	459 858	480 242	503 682	527 567
RMA Watanya	436 248	500 543	491 890	509 099	489 118	483 267
Wafa Assurance	129 267	145 301	301 899	349 318	383 459	414 383
Axa Assurance Maroc	471 943	496 691	395 090	384 689	389 992	394 890
Atlanta	219 387	222 672	243 718	207 719	326 481	346459
Sanad	111 255	118 808	122 145	136 622	116 235	137 528
Zurich	107 733	121 147	121 230	113 016	117 572	108 621
Marocaine Vie	24 831	19 980	18 821	16 727	18 370	22 128
MCMA	2 337	3 639	5 325	5 984	11 390	11 454
MAMDA	5 930	6 354	5 580	5 901	6 660	7 586
Marché	1 633 972	2 085 200	2 165 556	2 209 317	2 362 959	2 453 883

**Tableau 5-Chiffre d'affaire de la maladie des compagnies d'assurances (en mille DHS),
Source : Rapport Annuel 2013 de Saham Assurance**

Afin d'illustrer l'évolution du chiffre d'affaire de SAHAM Assurance par rapport à l'évolution du marché de la branche maladie, on a tracé la figure 4 :

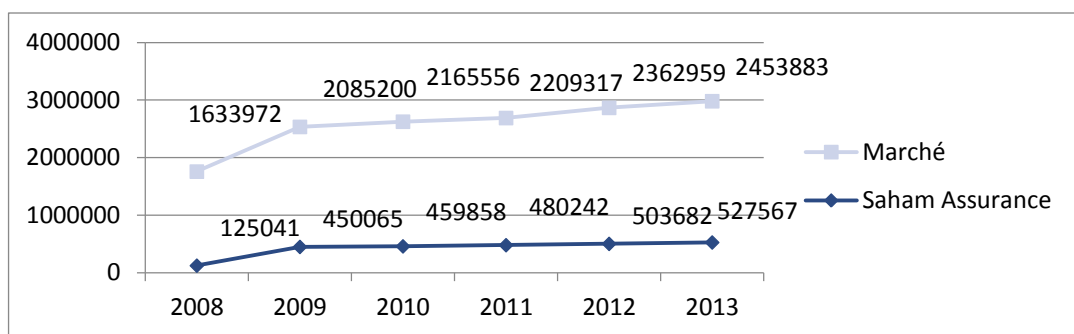


Figure 4- Evolution du chiffre d'affaire de l'assurance maladie et celui de SAHAM Assurance (en mille Dhs), Source : Rapport Annuel 2013 de Saham Assurance

Nous constatons que le chiffre d'affaire de Saham Assurance dans la branche Maladie ne cesse de s'accroître à l'instar de celui du Marché. Cependant une grosse évolution du chiffre d'affaire de Saham Assurance a été observée entre 2008 et 2009 et ceci peut s'expliquer par la fusion entre les deux assureurs CNIA et ES Saada.

II.2) Evolution part de marché assurance maladie :

Ce tableau contient la part de marché de chaque compagnie d'assurance pour la branche maladie :

	2008	2009	2010	2011	2012	2013
Saham Assurance	7,65%	21,58%	21,24%	21,74%	21,32%	21,50%
RMA Watanya	26,70%	24,00%	22,71%	23,04%	20,70%	19,69%
Wafa Assurance	7,91%	6,97%	13,94%	15,81%	16,23%	16,89%
Axa Assurance Maroc	28,88%	23,82%	18,24%	17,41%	16,50%	16,09%
Atlanta	13,43%	10,68%	11,25%	9,40%	13,82%	14,12%
Sanad	6,81%	5,70%	5,64%	6,18%	4,92%	5,60%
Zurich	6,59%	5,81%	5,60%	5,12%	4,98%	4,43%
Marocaine Vie	1,52%	0,96%	0,87%	0,76%	0,78%	0,90%
MCMA	0,14%	0,17%	0,25%	0,27%	0,48%	0,47%
MAMDA	0,36%	0,30%	0,26%	0,27%	0,28%	0,31%
Marché	100%	100%	100%	100%	100%	100%

Tableau 6-La part du marché de la maladie des différentes compagnies d'assurance, Source : Rapport Annuel 2013 de Saham Assurance

Nous constatons que Saham assurance détient la plus grande part de marché. La figure 5 illustre l'évolution de la part de marché de la compagnie de 2008 à 2013 :

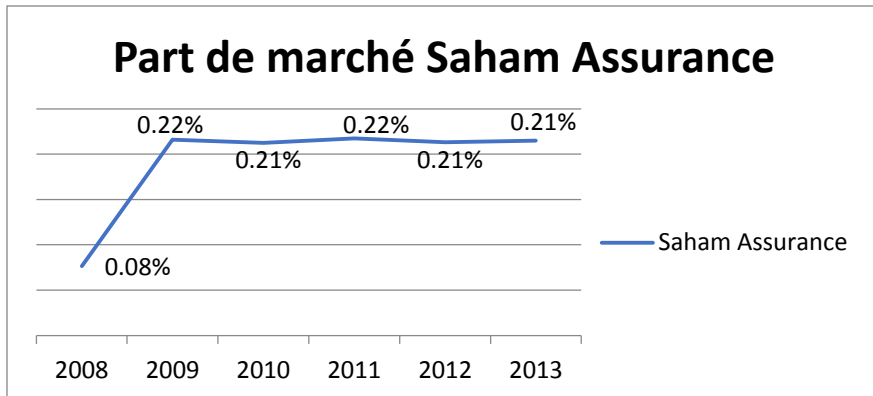


Figure 5-L'évolution de part de marché Saham Assurance

Nous remarquons que la part de marché de Saham Assurance reste au même niveau ce qui nous amène à déduire que d'un point de vue commercial la compagnie est fortement concurrentielle.

CHAPITRE 2 : PRESENTATION DES DONNEES DU PORTEFEUILLE

Présentation des données :

Les données utilisées se décomposent en 3 bases : la base « Affiliation » concernant les informations sur la population assurée, la base « prestations » portant sur l'ensemble des prestations médicales dont ont bénéficié ces mêmes assurés et la base « plafond » relative aux plafonds des prestations pour chaque groupe d'assurés.

Les données à notre disposition concernent les exercices s'étendant de 2011 à 2014.

I. La structure des bases utilisées :

1) La base « Affiliation » :

Cette base relate comme nous l'avons précisé précédemment les informations relatives à l'ensemble des assurés et bénéficiaires appartenant à un groupe donné.

Ainsi, nous retrouvons les variables suivantes :

- le numéro de Police relatif à l'entreprise contractante ;
- le numéro d'adhésion relatif à l'assuré principal au sein du groupe ainsi que des personnes à sa charge. Ce code est unique pour chaque famille au sein du groupe.
- le lien du bénéficiaire indiquant s'il s'agit d'un « assuré », « conjoint » ou « enfant ».
- la date de naissance du bénéficiaire.
- le sexe du bénéficiaire.
- la catégorie indiquant s'il s'agit d'un assuré actif ou retraité.
- la date d'affiliation au contrat.
- la date de sortie de ce contrat, prenant ainsi une valeur manquante si aucune date de sortie ou de résiliation n'est encore prévue.

2) La base « prestations » :

Pour chaque année d'exercice, Cette base comportait la consommation médicale par prestation pour chaque type de bénéficiaire au sein du groupe. Entre autres, nous retrouvons les variables suivantes jugées ainsi nécessaires à notre étude :

- le numéro de police
- le numéro d'adhésion
- le numéro de sinistre
- le lien
- Le montant des frais réels engagés pour la prestation médicale
- le montant remboursé par prestation
- le taux de remboursement
- le libellé de la prestation médicale
- l'état sinistre indiquant ainsi si le sinistre est « réglé », « rejeté » ou « en instance ».

3) La base « plafonds » :

Pour chaque numéro de police, Cette base contient les plafonds de la consommation pour chaque prestation plafonnée.

Ainsi, nous retrouvons les variables suivantes :

- Le Numéro de police
- Le plafond général
- Le plafond parodontie et prothèse dentaire (Assuré et conjoint)
- Le plafond monture
- Le plafond Odf, parodontie et prothèse dentaire (Enfant)
- Forfait naissance simple

Le plafond parodontie et prothèse dentaire concernant les assurés et le plafond Odf, parodontie et prothèse dentaire propre aux enfants vont être groupés sous un même plafond car ces deux couvertures sont de même nature.

II. La constitution des bases de données :

Les données fournies ont nécessité en premier lieu un traitement de manière à ne conserver que les données nécessaires à l'étude mais aussi à fiabiliser ces données. Ainsi, l'étude ne porte que sur les contrats collectifs. Les données concernant des contrats individuels ont donc été supprimées de la base. De même, les assurés principaux retraités ou inactifs ainsi que les personnes à leur charge ont été retirés de la base, vu que nous limiterons notre étude aux assurés actifs. Par ailleurs, la base « prestations » contenait pour le même individu au sein du groupe, plusieurs lignes indiquant chacune un sinistre qui a survenu durant cet exercice. Le but était de ne garder qu'une seule ligne par type de bénéficiaire pour chaque prestation sommant ainsi les nombres de sinistres et les frais engagés correspondants. Au final, nous récupérons une table qui contient pour chaque numéro de police référant au groupe, le nombre de sinistres survenus ainsi que la charge correspondante pour chaque type de bénéficiaire, par prestation médicale.

III. Regroupement des variables :

Dans les bases fournis par Saham assurance plusieurs prestations médicales existées. Dans un souci de simplification plusieurs d'entres elles ont été regroupées sous une même prestation globale. Ainsi ces nouvelles prestations feront l'objet de notre étude.

Ainsi, les différentes prestations à garder et leur nature (plafonnée ou non) sont les suivantes :

<i>Type de prestation</i>	<i>Nature</i>
Analyse (Biologie)	non plafonné
Dentaire	non plafonné
Hospitalisation	non plafonné
MONTURE	Plafonné
Maternité	Forfaitaire

NAISSANCE	Forfaitaire
Odf, parodontie et prothèse dentaire (Enfant)	Plafonné
Optique	non plafonné
Parodontie et prothèse dentaire (Assuré et Conjoint)	Plafonné
Pharmacie	non plafonné
Radiologie	non plafonné
Soins Ambulatoires	non plafonné

Tableau 7: Prestations offertes par Saham Assurance

IV. Création des variables :

Après ce premier traitement, un deuxième traitement a consisté à ajouter plusieurs variables explicatives à la table des affiliés, à partir des données existantes.

Ces variables sont :

- Les âges ont été calculés à partir des dates de naissance. Aucune anomalie n'a été rencontrée.
- L'exposition au risque qui correspond à la fraction d'années pendant laquelle l'assuré est sous risque durant une année d'exercice. L'année police est donc comprise entre 0 et 1 pour chaque exercice. Concrètement, pour l'exercice 2014, un assuré déjà couvert avant le 01/01/2014 et qui est toujours présent dans le portefeuille au 31/12/2014 aura une année police de 1.
- L'effectif du groupe ainsi que l'effectif de chaque type de bénéficiaire par groupe.

Pour terminer, une base de sinistre a été créée pour chacune des garanties, regroupant les différentes variables, les montants remboursés, la fréquence, le cout moyen ainsi que deux nouvelles variables :

- $\ln(\text{cout moyen})$ qui est la variable à expliquer dans le cas d'un modèle linéaire généralisé supposant une loi log-normale du cout moyen ;
- $\ln(\text{année police})$ qui est la variable offset dans le cas de la modélisation de la fréquence des sinistres avec une distribution poissonnienne ou binomiale négative avec une fonction de lien logarithmique.

Au final, et comme nous travaillons sur des contrats collectifs nous avons regroupé ces individus par numéro de police et par lien pour calculer ainsi l'âge moyen et l'exposition de chaque type de bénéficiaire au sein de chaque groupe.

CHAPITRE 3 : TARIFICATION A POSTERIORI

Introduction :

La tarification a posteriori a pour objectif la mise à jour des primes pures à partir de l'historique des sinistres des contrats. Cet objectif vise à obtenir un tarif corrigé et crédible.

L'objet de ce chapitre est de tarifier la consommation globale par police, nous allons arrêter un tarif global par police. Nous allons avancer une hypothèse forte qui sera vérifiée par la suite, cette dernière stipule que la différence des moyennes des consommations entre les différents types de prestation est non significative.

La tarification a posteriori repose sur une théorie dite de crédibilité. Cette dernière apparaît lors du passage du général au particulier, c'est-à-dire lors de la tarification d'un assuré pris isolément et non plus comme simple membre d'un groupe. Ainsi, la théorie de crédibilité fera l'objet de notre première tarification.

Dans ce chapitre, nous allons présenter la théorie de crédibilité. Ensuite nous allons nous attarder sur le modèle de Bühlmann qui constitue la base de cette théorie. Et le dernier point, traitera le modèle Bühlmann-Straub généralisation du premier modèle.

Ensuite un test d'égalité des moyennes sera effectué afin de vérifier l'hypothèse sur laquelle a reposé la tarification a posteriori.

I. La Théorie de crédibilité :

La théorie de crédibilité a été développée par les écoles suisse et scandinave, elle a pour objectif la mise à jour des primes pures à partir de l'historique des sinistres des contrats.

Intéressons-nous à une société d'assurance qui dispose d'un portefeuille de n contrats homogènes auxquels sont associées les variables aléatoires Y_{jt} pour $t = 1, \dots, T$ et $j = 1, \dots, n$ où Y_{jt} représente le montant annuel de sinistres du contrat j pour la $t^{\text{ème}}$ année.

Faisons l'hypothèse que pour chaque contrat j , les montants de sinistres annuels Y_{j1}, \dots, Y_{jT} sont indépendants et identiquement distribués.

Supposons en outre que ce portefeuille n'ait pas fait l'objet d'une segmentation a priori et qu'il soit demandé en $t=0$ le montant p_{coll} à tous les assurés. L'assureur peut être tenté de personnaliser les primes a posteriori en intégrant l'information qui lui est fournie par l'historique des sinistres des contrats. Cette mise à jour des tarifs s'effectue donc « au mérite » en récompensant les contrats les moins sinistrés, ce qui a pour conséquence de limiter l'aléa moral.

Considérons le contrat j et plaçons-nous en T . L'assureur dispose du passé (y_{j1}, \dots, y_{jT}) où

y_{jt} représente la charge annuelle des sinistres engendrés par la police j lors de la $t^{\text{ème}}$ année.

En moyenne, le contrat j a engendré une charge annuelle moyenne de sinistres de

$$\bar{p}_j = \frac{Y_{j1} + \dots + Y_{jT}}{T}$$

Réclamer \bar{p}_j à l'assuré du contrat j reviendrait à nier le principe de mutualisation qui est le fondement de l'activité d'assurance, aussi l'assureur peut avoir recours à un compromis qui le conduit à réclamer pour la période $[T, T + 1]$ la prime de crédibilité :

$$p_j = (1 - \alpha)p_{\text{coll}} + \alpha \bar{p}_j$$

Où α est le facteur de crédibilité, qui mesure la crédibilité que l'on peut accorder à la prime observée \bar{p}_j , c'est-à-dire aux sinistres passés de l'assuré, comme prédiction du niveau de risque futur.

Ainsi la problématique qui se pose à ce niveau est l'estimation du facteur de crédibilité.

Les modèles Bühlmann et Bühlmann-Straub qui seront présentés dans ce qui suit ont fait de ce facteur l'objet même de leur étude.

I.1) Présentation du modèle de Bühlmann :

Hans Bühlmann obtient avec des méthodes non paramétriques un facteur de crédibilité de la forme $\frac{A}{A+B}$. En effet pour aboutir à ce résultat, cet actuaire a forcé la prime à être linéaire. Cette solution a constitué une innovation dans la théorie de crédibilité.

1) Les hypothèses du modèle :

Ce modèle simpliste repose sur les hypothèses suivantes :

- H1 : Chaque contrat est caractérisé par un vecteur aléatoire $(R_j, Y_{j1}, \dots, Y_{jT})$ où :
 - Y_{jt} est le montant de sinistre relatif au contrat i pour l'année t .
 - R_j est un paramètre de risque inconnu.

Les vecteurs $(R_j, Y_{j1}, \dots, Y_{jn})$ sont indépendants pour $j = 1, 2, \dots, n$.

- H2 : Les variables R_1, \dots, R_n sont indépendants et identiquement distribués.
- H3 : Conditionnellement à R_j , les variables aléatoires Y_{jt} , $t = 1, 2, \dots, T$ sont indépendantes et identiquement distribuées.
- H4 : Il suppose que :

$$E[Y_{j1} / R_j] = \mu(R_j) \quad \text{et} \quad \text{Var}[Y_{j1} / R_j] = \sigma^2(R_j)$$

2) Estimation du facteur de crédibilité :

Afin de simplifier les formules les formules suivantes seront adoptées :

- ✓ $M^2 = \text{var}(\mu(R_j))$: mesure la part de la variance due à l'hétérogénéité du portefeuille.
- ✓ $\gamma^2 = E[\sigma^2(R_j)]$: mesure la part du hasard dans la variance a priori.
- ✓ $\mu = E[\mu(R_j)]$

Dans la suite nous aurons besoin des deux résultats suivants :

- $\text{Cov}[Y_{jt}, Y_{js}] = \delta_{s,t} \gamma^2 + M^2$
- $\text{Cov}[Y_{jt}, \mu(j)] = M^2$

Démonstration :

Démontrons la première égalité, en effet on a :

$$\begin{aligned} \text{Cov}[Y_{jt}, Y_{js}] &= E[\text{Cov}[Y_{jt}, Y_{js} / R_j]] + \text{Cov}[E[Y_{jt} / R_j], E[Y_{js} / R_j]] \\ &= \delta_{s,t} E[\sigma^2(R_j)] + \text{var}(\mu(R_j)) \end{aligned}$$

Car les Y_{jt} et Y_{js} sont indépendants conditionnellement à R_j .

A présent démontrons la deuxième égalité :

$$\begin{aligned} \text{On a } \text{Cov}[Y_{jt}, \mu(R_j)] &= E[\text{Cov}[Y_{jt}, \mu(R_j) / R_j]] + \text{Cov}[E[Y_{jt} / R_j], E[\mu(R_j) / R_j]] \\ &= M^2 \end{aligned}$$

Puisque conditionnellement à R_j , $\mu(R_j)$ est une constante.

3) Prime de crédibilité :

Le modèle de crédibilité de Bühlmann utilise cette approximation linéaire puisqu'il propose de retenir pour la période $[T, T + 1]$ une prime de la forme :

$$p(Y_{j1}, \dots, Y_{jT}) = c_{j0} + \sum_1^T c_{jt} Y_{jt} ,$$

Où les c_t sont choisis de manière à minimiser l'écart quadratique moyen

$$E[\{\mu(R_j) - c_{j0} - \sum_1^T c_{jt} Y_{jt}\}^2]$$

Ce modèle repose sur les hypothèses formulées ci-haut.

En annulant les dérivées partielles par rapport à c_{jt} , $t = 1, \dots, T$ on aboutit aux systèmes d'équations suivants :

$$\begin{cases} E\{\mu(R_j) - c_{j0} - \sum_1^T c_{jt} Y_{jt}\} = 0 \\ E[Y_{jk} \{\mu(R_j) - c_{j0} - \sum_1^T c_{jt} Y_{jt}\}] = 0, \quad k = 1, 2, \dots, T \end{cases}$$

Les deux équations nous donnent les résultats suivants :

$$\begin{cases} \mu(1 - \sum_1^T c_{jt}) - c_{j0} = 0 \\ \text{Cov}[Y_{jk}, \mu(R_j)] - \sum_1^T c_{jt} \text{Cov}[Y_{jt}, Y_{js}] = 0, \quad k = 1, 2, \dots, T \end{cases}$$

D'après le lemme on retrouve le résultat suivant :

$$\begin{aligned} M^2 - \sum_{t=1}^T c_{jt} (\delta_{t,k} \gamma^2 + M^2) &= 0 \\ \Leftrightarrow M^2 (1 - \sum_{t=1}^T c_{jt}) &= c_{jk} \gamma^2, \quad k = 1, 2, \dots, T. \end{aligned}$$

En sommant sur k les n relations ci-dessus, on aboutit à l'équation suivante :

$$M^2 (1 - \sum_{t=1}^T c_{jt}) = \sum_{k=1}^T c_{jk} \gamma^2$$

Donc on a :
$$\sum_{t=1}^T c_{jt} = \frac{M^2}{\gamma^2 + M^2},$$

D'où :
$$c_{jk} = \frac{M^2}{\gamma^2} \left(1 - \frac{M^2}{\gamma^2 + M^2}\right) = \frac{M^2}{\gamma^2 + M^2}, k = 1, 2, \dots, T.$$

Et
$$c_{i0} = \mu \frac{M^2}{\gamma^2 + M^2}$$

Finalement l'estimateur de crédibilité de $\mu(R_j)$ est donné par l'équation suivante :

$$\mu \frac{M^2}{\gamma^2 + M^2} + \frac{M^2}{\gamma^2 + M^2} = \sum_{t=1}^T Y_{jt}$$

Cette prime représente la prime de crédibilité pour la police j avec un facteur de crédibilité α_j , où :

$$\alpha_j = \frac{TM^2}{\gamma^2 + TM^2}$$

Le modèle de Bühlmann estime les paramètres μ , γ^2 et M^2 par les expressions suivantes :

$$\left\{ \begin{array}{l} \hat{\mu} = \sum_{j=1}^n \sum_{t=1}^T Y_{jt} , \\ \widehat{\gamma^2} = \frac{1}{n} \sum_{j=1}^n \widehat{\sigma_j^2} , \\ \widehat{M^2} = \sum_{j=1}^n (\bar{Y}_j - \bar{Y}_{..})^2 - (n-1) \widehat{\gamma^2} \end{array} \right.$$

Avec :
$$\widehat{\sigma_j^2} = \frac{1}{T-1} \sum_{t=1}^T (Y_{jt} - \bar{Y}_j)^2$$

I.2) Présentation du modèle de Bühlmann –Straub :

Le modèle de Bühlmann-Straub est une extension du modèle de Bühlmann qui intègre des pondérations sur les observations. En effet dans le précédent modèle repose sur l'hypothèse que chaque sinistre aura le même poids dans la détermination de la prime de crédibilité. Pour corriger cette hypothèse peu réaliste, Le modèle de Bühlmann-Straub prend en compte le fait que les classes de contrats sont souvent d'importances inégales, notamment par le nombre de contrats qu'elles regroupent.

Ce modèle repose presque sur les mêmes hypothèses, seulement la quatrième qui devient :

$$E[Y_{j1} / R_j] = \mu(R_j) \quad \text{et} \quad \text{Var}[Y_{j1} / R_j] = \frac{1}{w_{jt}} \sigma^2(R_j)$$

Avec w_{jt} le poids traduisant l'importance de la police i au cours de l'année t au sein de notre portefeuille.

L'estimateur de crédibilité de $\mu(R_j)$ devient :

$$\mu \frac{M^2}{\gamma^2 + w_j M^2} + \frac{w_j M^2}{\gamma^2 + w_j M^2} = \frac{1}{w_j} \sum_{k=1}^T w_{jk} Y_{jt}$$

Le facteur de crédibilité pour la police j devient :

$$\alpha_j = \frac{w_j M^2}{\gamma^2 + w_j M^2}$$

Les paramètres de structure du modèle deviennent :

$$\left\{ \begin{array}{l} \hat{\mu} = \frac{1}{w_{..}} \sum_{j=1}^n \sum_{t=1}^T w_{jt} Y_{jt} , \\ \hat{\gamma}^2 = \frac{1}{n} \sum_{j=1}^n \hat{\sigma}_j^2 , \\ \hat{M}^2 = \frac{1}{w_{..}} \sum_{j=1}^n w_{j.}^2 (\bar{Y}_{j.} - \bar{Y}_{..})^2 - (n-1) \hat{\gamma}^2 \end{array} \right.$$

Avec : $\hat{\sigma}_j^2 = \frac{1}{T-1} \sum_{t=1}^T w_{jt} (Y_{jt} - \bar{Y}_{j.})^2$.

I.3) Application de la théorie de crédibilité sur la population étudiée :

1) Présentation des données utilisées :

En utilisant comme variable de fusion le numéro de la police nous avons calculé la charge totale annuelle par police pour la période s'étalant de 2011 à 2014. Le choix de cette variable de fusion se justifie par le caractère fixe de cette variable à travers les années, nous avons retenues 451 entreprises. La nouvelle table contient les variables suivantes :

- ✓ La police
- ✓ L'effectif de la police
- ✓ Le plafond général de la police
- ✓ Les charges totales par année, par plafond général et par effectif.

2) Les classes retenues en crédibilité :

Afin de segmenter notre portefeuille Nous avons retenues deux variables jugées les plus discriminantes à savoir la taille de l'entreprise et le plafond général de chaque police car ces deux variables sont communes à tout le groupe.

Ainsi la segmentation s'effectuera sur l'observation des quantiles. Les variables taille de l'effectif et plafond général ont respectivement 216 et 22 modalités, donc la variable effectif

du groupe sera segmenté en quatre classes et la variable plafond général sera segmenté en 3 classes.

Le tableau 8 représente les quantiles de la variable Effectif du groupe :

Effectif du groupe	
Quantiles	Valeur
25%	30
50%	70
75%	200

Tableau 8: Classes de la variable Effectif

Ainsi on retiendra les 4 classes présentes dans le tableau précédent. De même la segmentation de la variable Plafond Général est présentée dans le tableau 9.

Plafond Général	
Quantile	Valeur
33%	35 000
66%	50 000

Tableau 9: Classes de la variable Plafond Général

Ainsi les classes retenues sont retracés dans le tableau 10, ce tableau nous renseigne aussi sur l'effectif de chaque classe :

Classe	Nombre d'entreprise	Effectif	Plafond Général
1	142	<=30	<=35000
2	72	<=30] 35000-50000]
3	15	<=30	>50000
4	51] 30-70]	<=35000
5	23] 30-70]] 35000-50000]
6	9] 30-70]	>50000
7	30] 70-200]	<=35000
8	34] 70-200]] 35000-50000]
9	7] 70-200]	>50000
10	21	>200	<=35000
11	24	>200] 35000-50000]
12	23	>200	>50000

Tableau 10-les classes de Tarification à posteriori

Nous avons étudié la consommation moyenne par assurées par classe sur les quatre années d'études, cette étude est retracée sur le tableau 11 :

Classe	Consommation moyenne 2011	Consommation moyenne 2012	Consommation moyenne 2013	Consommation moyenne 2014
1	980.86	1 086.78	1 161.78	1 015.67
2	1 770.71	1 770.76	2 030.17	2 056.87
3	1 468.18	1 373.61	1 273.32	1 179.53
4	721.83	822.38	839.06	763.47
5	1 574.26	1 548.18	1 553.92	1 651.36
6	987.17	1 319.67	1 147.02	999.62
7	585.99	625.41	644.50	611.95
8	1 286.01	1 157.62	1 135.46	1 079.38
9	1 638.21	1 462.60	1 273.60	1 185.39
10	362.59	449.29	599.07	339.19
11	539.66	490.20	465.61	278.55
12	357.74	373.62	427.99	253.45

Tableau 11-La consommation moyenne par classe sur les quatre années d'études (en Dhs)

D'après le tableau précédent, nous remarquons que la consommation moyenne par assuré varie d'une classe à une autre. Nous constatons aussi que la consommation diminue avec l'effectif ce qui nous amène à conclure que plus la taille de l'entreprise est grande plus la consommation moyenne par assuré diminue.

3) Estimation des paramètres et résultats du modèle de Bühlmann :

✓ Estimation des paramètres

Un programme VBA a été conçu pour calculer les formules présentés dans le chapitre précédent. Les résultats obtenus pour les différentes classes sont présentés dans le tableau suivant :

Classe	γ^2	M^2	α
1	553 459,20	338 755,20	0,71
2	187 113,50	83 161,56	0,64
3	637 509,30	1 611 481,84	0,91
4	584 816,70	623 291,48	0,81
5	604 494,20	430 120,87	0,74
6	553 586,40	294 092,78	0,68

7	598 981,50	206 791,23	0,58
8	480 378,90	1 595 544,20	0,93
9	363 761,70	322 425,14	0,78
10	282 917,00	157 429,62	0,69
11	184 398,90	131 206,91	0,74
12	586 173,90	551 282,60	0,79

Tableau 12- les paramètres de structure du modèle de Bühlmann

Nous constatons que pour la 3^{ème} et 8^{ème} classe le facteur de crédibilité est supérieur à 90% ainsi la prime de crédibilité que devront payer les entreprises ayant ces caractéristiques sera à peu près égale à leur prime historique.

✓ Exemple des résultats de la 5^{ème} classe

A présent nous avons toutes les informations pour calculer la prime de crédibilité. De ce fait, nous avons choisi, pour un alpha donné (classe 5), quelques exemples de calcul de cette prime. Cet exemple sera présenté au tableau 13.

La prime collective est calculée par la relation suivante :

$$N * \text{Fréquence} * \text{Coût Moyen}$$

Avec :

- N : L'effectif Globale à savoir le nombre des assurées, des conjoints et des enfants
- Fréquence : représente la fréquence des sinistres de chaque classe, cette fréquence est calculée par :

$$\frac{\text{Nombre de Sinistre annuelle de 2014 par classe}}{\text{Exposition au risque}}$$

- Coût Moyen : représente le coût moyen des sinistres de cette classe, ce coût moyen est calculé par :

$$\frac{\text{Charge totale annuelle de 2014 par classe}}{\text{Exposition au risque}}$$

Police	Prime historique(en Dhs)	Alpha	Prime collective (a priori)(en Dhs)	Prime de crédibilité(en Dhs)
1	71 068,14	0,74	24 414,55	58 938,20
2	27 292,32	0,74	24 414,55	26 544,10
3	10 761,47	0,74	24 414,55	14 311,27
4	24 011,51	0,74	24 414,55	24 116,30
5	20 548,20	0,74	24 414,55	21 553,45
6	38 650,52	0,74	24 414,55	34 949,17

Tableau 13-Exemples de prime de crédibilité pour la 5^{ème} classe

Trois cas peuvent se présenter :

- Les entreprises ayant une prime a priori chère : sont les entreprises ayant une prime historique faible. De ce fait, la prime de crédibilité à accorder à ces entreprises est moins que leur prime a priori.
- Les entreprises ayant un tarif a priori sous-estimé : sont les entreprises ayant des primes de crédibilité supérieures à celles effectivement réclamées, ce qui signifie que leur historique n’a pas joué en leur faveur.
- Les entreprises ayant le même tarif (a posteriori et a priori) : sont généralement les entreprises ayant un comportement de consommation stable dans le temps.

4) Estimation des paramètres du modèle de Bühlmann-Straub :

Comme cela a été présenté ci-haut le modèle de Bühlmann-Straub est une généralisation du modèle simple de Bühlmann avec l’introduction d’un poids à chaque observation. En effet le poids retenu dans notre étude est l’effectif de l’entreprise, à cet effet une entreprise comptant par exemple 100 individus est beaucoup plus exposée au risque qu’une entreprise avec 10 individus.

Après introduction de l’effectif de l’entreprise comme poids, l’estimation du modèle devient :

Classe	γ^2	M^2	α
1	14 773 329,51	33 827 990,82	0,86
2	6 724 750,81	31 415 743,54	0,74
3	4 811 403,24	38 502 657,46	0,94
4	2 134 545,75	31 141 867,78	0,81
5	5 133 846,62	35 392 605,27	0,88
6	4 005 505,17	38 725 073,71	0,99
7	9 908 334,59	15 206 181,86	0,84
8	3 739 288,06	27 812 129,53	0,89
9	4 192 721,81	19 630 814,74	0,98
10	8 945 050,37	35 920 388,39	0,91
11	10 061 805,20	25 145 691,12	0,87
12	10 410 472,72	38 353 929,93	0,77

Tableau 14-Les paramètres de structure du modèle Bühlmann-Straub

Nous remarquons que pour certaine classe le modèle de Bühlmann-Straub élimine toute mutualisation entre les bénéficiaires car les facteurs de crédibilité calculé sont proches de 1, du coup la prime de crédibilité sera à peu près égal à la prime historique. Ainsi les entreprises avec une forte consommation devront payer un tarif élevé et celles avec une faible consommation devront payer une prime faible. Ce résultat peut aussi s’expliquer par le nombre réduit d’entreprises dans ces classes, ainsi l’information n’est pas disponible pour bien estimer nos paramètres.

Le principe de l'assurance se base sur la mutualisation, hors on observe qu'avec ce modèle cette dernière disparaît ou presque, donc les primes de crédibilité ne seront pas calculés pour ce modèle car les résultats obtenus sont contradictoires avec le principe même de l'assurance.

Conclusion :

A travers cette partie nous avons présenté le cadre théorique de l'application de la théorie de crédibilité qui reste un atout majeur dans la tarification de la consommation globale par contrat.

Dans notre cas l'application de cette méthode représente plusieurs limites à savoir :

- ✓ Cette théorie nécessite un grand horizon de temps afin d'aboutir à des résultats pertinents et reflétant le risque couvert.
- ✓ L'échantillon sur lequel l'étude a été réalisée ne représente que 22,27% de l'ensemble du portefeuille, ce qui nous laisse sceptique quant à la précision et la robustesse des résultats apportés par cette méthode.

Pour conclure, dans cette partie nous avons essayé surtout de montrer l'apport théorique que peut offrir une tarification en utilisant l'historique des contrats.

L'hypothèse avancée ci-haut dans l'introduction du chapitre sera vérifiée dans la partie suivante afin de confirmer ou entériner la démarche retenue dans ce chapitre à savoir la tarification de la charge annuelle toute prestation confondue.

II. Analyse de la variance :

Dans cette partie nous allons effectuer une étude détaillée sur la consommation moyenne des différentes prestations par type de bénéficiaire afin de confirmer ou entériner l'hypothèse avancée au début du chapitre. En effet nous essaierons à travers l'étude d'égalité des moyennes de vérifier si la variable représentant la charge par prestation présente le même comportement vis-à-vis de la variable type de prestation.

Cette étude sera réalisée par type de bénéficiaire, les résultats présentés ci-dessous ne concernent que les assurées, pour les conjoints et les enfants les résultats sont joints en annexe 3.

II.1) Test d'égalité des variances :

Pour appliquer une ANOVA plusieurs hypothèses doivent être vérifiées parmi lesquelles l'homoscédasticité ou l'égalité des variances. Pour cela on fait appel au test de Levene afin de tester les hypothèses suivantes :

$$H_0 : \text{Les variances sont égales} \quad \text{Vs} \quad H_1 : \exists (i, i') \text{ tel que } \sigma_i^2 \neq \sigma_{i'}^2$$

✓ Test de Levene :

Ce test réalise une analyse de variance sur les écarts absolus des valeurs aux moyennes des groupes respectifs. Si le test de Levene est statistiquement significatif alors l'hypothèse d'homogénéité des variances doit être rejetée.

La statistique du test s'écrit sous la forme suivante :

$$W = \frac{(N - k)}{(k - 1)} \frac{\sum_{i=1}^k N_i (Z_{i.} - Z_{..})^2}{\sum_{i=1}^k \sum_{j=1}^{N_i} (Z_{ij} - Z_{i.})^2}$$

Avec N : taille de l'échantillon

N_i : Effectif du groupe i

k : nombre de groupes

Z_{ij} : $| y_{ij} - \tilde{y}_i |$ où \tilde{y}_i est la médiane du groupe j et y_{ij} est une l'observation i du groupe j

$Z_{i.}$: la moyenne des Z_{ij} pour chaque groupe j

W suit approximativement une loi $F(\alpha, k - 1, N - k)$.

Le tableau 15 nous résume le résultat obtenu après application du test Levene au seuil de 5% :

Statistique de Levene	ddl1	ddl2	P-value
35.185	11	10983	.000

Tableau 15-Statistique de Levene

Nous constatons que la p-value est inférieur au seuil de significativité donc l'hypothèse d'égalité des variances sera rejetée, ainsi l'analyse de variance à un seul facteur ne peut être utilisé puisque une condition n'est pas respectée.

Afin de comparer la moyenne des consommations des différentes prestations on aura recours à des tests non paramétriques qui seront détaillés dans ce qui suit :

II.2) Test d'égalité des moyennes :

1) Test de Welch :

Le test t de Welch est une adaptation du test t de student permettant de comparer des groupes d'échantillons lorsque les variances sont différentes.

Le test confronte les deux hypothèses suivantes :

$$'H_0: m_A = m_B ' \text{ Vs } 'H_1: m_A \neq m_B '$$

Avec m_A et m_B représentent respectivement la moyenne du groupe A et celle du groupe B. La statistique du test est calculée par la formule suivante :

$$t = \frac{m_A - m_B}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}}$$

Avec : A et B sont les deux groupes différents à comparer.

m_A et m_B représentent respectivement la moyenne du groupe A et celle du groupe B.

n_A et n_B représentent respectivement la taille du groupe A et celle du groupe B.

S_A et S_B représentent respectivement l'écart type du groupe A et celui du groupe B

NB : Contrairement au test de student classique, la formule du test t de Welch fait intervenir la variance calculée de chacun des deux groupes (S_A^2 et S_B^2). En d'autres termes on n'utilise pas la variance commune calculée sur les deux groupes en même temps.

Le calcul du degré de liberté du test t de Welch se fait selon la formule suivante :

$$\vartheta = \left(\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B} \right)^2 / \left(\frac{S_A^4}{n_A^2(n_A-1)} + \frac{S_B^4}{n_B^2(n_B-1)} \right)$$

La statistique de Welch suit approximativement une loi de student de degré de liberté ϑ et de seuil $\alpha=5\%$.

2) Résultats du test :

Test	Statistique	Ddl1	Dd2	P-value
Welch	17.50	11	1110.54	0.00

Tableau 16-Statistique de Welch

D'après la P-value du test l'hypothèse nulle d'égalité des moyennes sera rejetée.

Cette décision nous indique que la consommation diffère par type de prestation, ainsi nous concluons que des différences significatives existent dans le comportement des assurées vis-à-vis de la couverture qu'ils ont choisie.

Nous allons détailler encore davantage notre analyse de la consommation inter prestation afin de pouvoir détecter la prévenance de cet écart. A cet effet nous allons effectuer une comparaison multiple de la moyenne de différentes prestations.

Pour Effectuer cette analyse plus approfondie un test Post-Hoc s'impose. Puisque l'hypothèse d'égalité des variances a été rejetée donc le test de Bonferroni, Scheffé et Tukey ne sont plus applicables, on fera appel au test T2 de Tamhane pour calculer la différence moyenne de la consommation entre les différentes prestations.

Pour ce test de comparaison multiple des moyennes, deux moyennes sont significativement différentes si :

$$|\bar{y}_i - \bar{y}_j| \geq Q_{ij}^* R_{\varepsilon,r,\vartheta}$$

Avec :

- $Q_{ij}^* = \sqrt{\frac{S_i^2}{n_i} + \frac{S_j^2}{n_j}}$
- $R_{\varepsilon,r,\vartheta} = \sqrt{F_{\gamma,1,\vartheta}} = t_{\gamma,\vartheta}$ où $\gamma = 1 - (1 - \varepsilon)^{1/k^*}$
- ϑ représente le degré de liberté, sa formule a été définie ci-haut dans la page
- ε Experiment wise error rate under the complete null hypothesis
- $k^* = k(k-1)/2$ avec k nombre de groupe
- r Number of steps between means

3) Résultats pour la prestation ‘Analyse (Biologie)’ :

Prestation		Différence de moyennes (I-J)	Erreur standard	P-value	Intervalle de confiance à 95%	
I	J				Borne inférieure	Borne supérieure
Analyse (Biologie)	Dentaire	-1634	2238	1	-9167	5900
	Hospitalisation	-39990	5999	0	-60230	-19751
	Monture	4977	1494	0,056	-52	10006
	Maternité	1861	2137	1	-5346	9068
	Naissance	4184	1574	0,411	-1124	9491
	Optique	-942	1767	1	-6889	5004
	Pharmacie	-18355	3103	0	-28800	-7911
	Radiologie	3024	1466	0,929	-1909	7957
	Soins Ambulatoires	-9718	2494	0,007	-18114	-1324
	Parodontie et prothèse dentaire	-17526	3475	0	-29237	-5816

Tableau 17- Résultat du test d'égalité de la consommation moyenne par rapport à la prestation « Analyse (biologie) »

D'après les p-value nous constatons qu'il existe une différence significative de le moyenne de consommation entre les prestations ‘Analyse (Biologie)’ et les prestations ‘Hospitalisation’, ‘Pharmacie’, ‘Soins Ambulatoires’ et ‘Parodontie et prothèses dentaires’, d'ailleurs nous constatons que l'intervalle de confiance de la différence entre la consommation moyenne de la prestation ‘Analyse (Biologie)’ et les prestations citées auparavant ne contiennent pas la valeur 0 ce qui nous amène aux mêmes conclusions.

En ce qui concerne les autres prestations, les résultats par type de bénéficiaire sont présentés à l'annexe 3.

Conclusion :

Le test d'égalité de la moyenne de consommation par type de prestation que nous avons effectué respectivement sur les assurés, les conjoints puis les enfants, dégage une différence significative de comportement en termes de consommation d'une prestation à une autre. Ainsi la démarche de la tarification de l'ensemble de la police toute prestation confondue qu'on a adoptée au chapitre précédent n'est pas légitime. En effet cette différence de consommation moyenne entre les prestations nous conduit à créer un outil de tarification par poste et non pas par charge totale. A cet effet chaque prestation sera tarifiée à part par le modèle GLM.

CHAPITRE 4 : TARIFICATION A PRIORI

Introduction :

Après avoir prouvé la non significativité de l'hypothèse d'égalité des moyennes entre les prestations une tarification par poste s'impose, ainsi d'autres variables propres aux prestations seront introduites et par cela une analyse de données de ce portefeuille d'assurés s'impose afin d'identifier les corrélations qui peuvent exister entre les variables explicatives retenues.

Vu les résultats obtenues et les conclusions émises aux chapitres X, nous avons été contraint de prospecter de nouvelle méthode de tarification. Ainsi L'objectif de ce chapitre est de définir une nouvelle tarification adéquate des contrats collectifs en assurance maladie. La méthode qui sera utilisée est l'approche « Fréquence * Coût moyen » en modélisant les deux indicateurs par les modèles linéaires généralisés.

Pour ce faire nous allons présenter le cadre théorique que nous allons adopter et mettre en pratique l'approche théorique adoptée pour l'élaboration du modèle le plus adéquat. Ainsi, Nous allons tout d'abord élaborer un ensemble de statistiques descriptives permettant d'analyser le portefeuille étudié, réaliser une segmentation des variables tarifaires pour enfin mettre en pratique l'approche adopté pour déterminer le tarif global du groupe par type de bénéficiaires.

Les données sur lesquels porteront l'ensemble de l'étude statistique suivante sont relatifs aux années 2013-2014.

I. Taux de remboursement pondéré :

Afin d'adapter au mieux le tarif aux différents comportements de consommation qui peuvent exister, une prime est calculée pour chaque type de prestation médicale. Ainsi, un regroupement des prestations médicales nous a été confié par la compagnie afin de réduire le nombre de prestations à tarifier. Cette étape consiste à regrouper les prestations par groupes de prestations. Toutefois, chaque prestation a un taux de remboursement différent des autres. Par la suite, le taux de remboursement équivalent à la prestation regroupée est obtenu en calculant la moyenne pondérée des taux de remboursement de chaque prestation par la consommation des bénéficiaires pour cette même prestation. Au final, Nous obtenons ainsi pour chaque type de bénéficiaire au sein du groupe, le nombre total de sinistres et le montant total remboursé pour chacun des groupes de prestation sur la période d'observation.

II. L'Analyse en composantes principales :

L'Analyse en Composantes principales (ACP) fait partie du groupe des méthodes descriptives multidimensionnelles appelées méthodes factorielles. Dans la mesure où ce sont des méthodes descriptives, elles ne s'appuient pas sur un modèle probabiliste, mais elles dépendent d'un modèle géométrique. L'ACP propose, à partir d'un tableau rectangulaire de données comportant les valeurs de p variables quantitatives pour n unités (appelées aussi individus), des représentations géométriques de ces unités et de ces variables. Ces données peuvent être issues d'une procédure d'échantillonnage ou bien de l'observation d'une population toute entière. Les représentations des unités permettent de voir s'il existe une structure, non connue a priori, sur cet ensemble d'unités. De façon analogue, les représentations des variables permettent d'étudier les structures de liaisons linéaires sur l'ensemble des variables considérées. Ainsi, on cherchera si l'on peut distinguer des groupes

dans l'ensemble des unités en regardant quelles sont les unités qui se ressemblent, celles qui se distinguent des autres, etc. Pour les variables, on cherchera quelles sont celles qui sont très corrélées entre elles, celles qui, au contraire ne sont pas corrélées aux autres, etc.

1) Présentation des variables :

Les variables sur lesquelles l'analyse en composantes principale sera effectuée sont :

- L'effectif de l'entreprise (Assurés)
- Le plafond général
- Le taux de remboursement
- Le plafond monture
- Le plafond parodontie et prothèse dentaire (plus la prestation Odf pour les enfants)
- Le forfait Naissance simple

Les variables suivantes Effectif des conjoints et des enfants non pas été introduites pour des raisons stratégiques et commerciaux. En effet SAHAM Assurance cherche à promouvoir davantage sont produits ainsi en se référant au point de vue avancé par la compagnie une entreprise ne doit en aucun cas payer plus ou moins parce que les ayants droits sont nombreux ou pas ainsi ces deux variables tarifaires seront écartées des modèles ajustées.

Après divers débats avec le manager responsable de la branche Maladie à SAHAM Assurance, nous avons conclu que la variable Age moyen de l'entreprise ne va pas être retenu dans la tarification car cette variable est utilisée dans la tarification individuelle.

2) Réduction des dimensions :

Le tableau 18 représente les valeurs propres des facteurs ainsi que l'inertie propre à chaque axe :

Variance totale expliquée						
Composante	Valeurs propres initiales			Somme des carrés des facteurs retenus pour la rotation		
	Total	% de la variance	% cumulés	Total	% de la variance	% cumulés
1	2,602	43,373	43,373	2,575	42,911	42,911
2	1,079	17,983	61,356	1,107	18,445	61,356
3	,902	15,034	76,390			
4	,767	12,781	89,171			
5	,460	7,673	96,844			
6	,189	3,156	100,000			

Tableau 18: L'inertie expliquée par chaque axe factorielle

Nous retenons les axes dont la valeur propre est supérieur ou égale à 1 donc d'après le tableau des valeurs propres des axes nous retenons les deux premiers axes.

Le premier axe explique à lui seul 42,91% de l'inertie totale alors que le deuxième axe représente 18,44% de l'information totale. Ces deux premiers axes expliquent 61,356% de l'inertie totale. D'après les résultats de la rotation varimax nous concluons aussi que nous allons retenir deux axes expliquant 61,356% de l'information totale.

3) La qualité de représentation des variables sur les deux axes retenues :

Le tableau 19 représente la qualité de représentation ou le score de chaque variable sur les deux axes retenus :

Matrice des composantes après rotation ^a		
Variable	Composante	
	1	2
Taux de remboursement	,550	-,372
Plafond Général	,369	,474
Forfait naissance simple	,872	,179
Plafond monture	,888	,190
Le plafond parodontie et prothèse	,763	-,050
Effectif de l'entreprise	-,066	,820

Tableau 19: Le score de chaque variable sur les axes factoriels retenus

4) Interprétations des axes :

D'après la qualité de représentation de chaque variable sur les deux axes nous constatons que :

- Le premier axe représente les caractéristiques propres à chaque prestation au sein du même groupe.
- Le deuxième axe représente les caractéristiques communes à tout le groupe à savoir l'effectif du groupe et le plafond-général de la couverture maladie.

Le figure 6 nous confirme les interprétations émises pour chaque axe car on remarque que le premier axe représente les sous-plafonds ainsi que le taux de remboursement et le deuxième axe représente l'effectif et le plafond-Général ce qui confirme d'avantage nos interprétations.

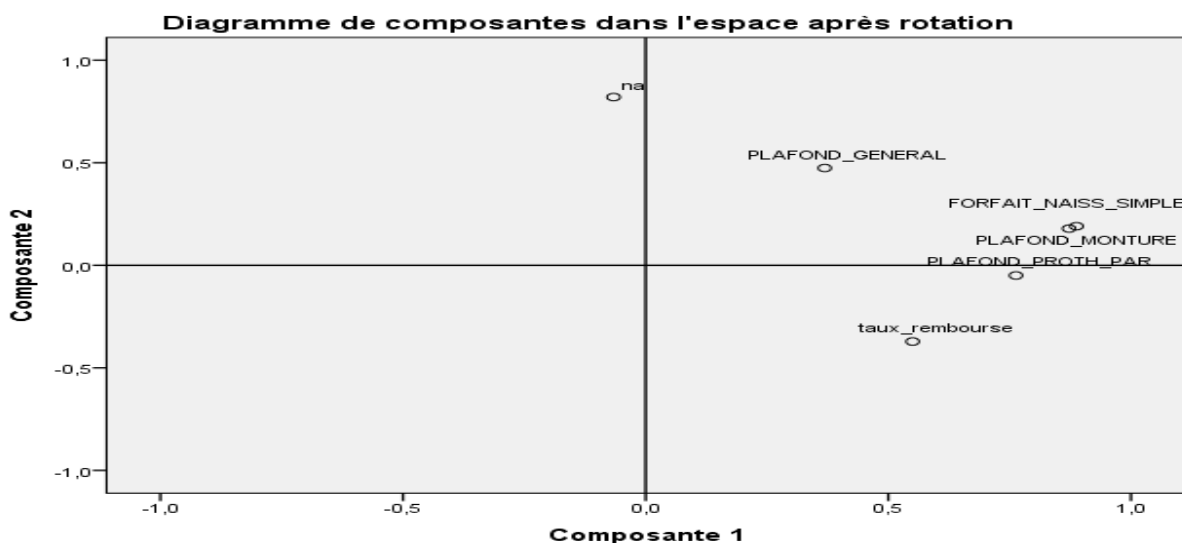


Figure 6-Diagramme des composantes dans l'espace

III. Segmentation et codification des variables :

La segmentation des variables est une étape importante du processus de modélisation. En effet elle nous permet de passer d'une infinité de classes de risques à un nombre fini de classes de risque facilement interprétables. Pour faire la segmentation nous allons utiliser les arbres de décision. Il existe plusieurs algorithmes de construction de ces arbres : l'algorithme CRT, l'algorithme CHAID etc.... Nous avons retenue l'algorithme CHAID pour notre segmentation.

III.1) Présentation de l'algorithme CHAID :

Chaid est l'acronyme de 'Chi-square Automatic Interaction Detected', il désigne un algorithme de segmentation de variable. Il se base pour cela sur une variable dépendante qui lui servira de guide. L'algorithme subdivise la population initiale en plusieurs ensembles distincts en se basant sur les variables indépendantes les plus significatives, chacun des groupes ainsi obtenu est aussitôt subdivisé en deux sous-groupes distincts ou plus en considérant les variables restantes. On répète ce processus jusqu'à la fin, soit toutes les variables sont utilisées, soit une condition spécifiée par l'utilisateur est réalisée. Le résumé de toutes ces démarches est donné sous forme d'un arbre de décision où chaque nœud représente un regroupement d'observations. Il y a trois étapes importantes lors de l'exécution de l'algorithme :

1) Préparation des prédicteurs :

Tout d'abord, l'algorithme va créer des prédicteurs catégoriels à partir de chacun des prédicteurs continus, en répartissant la distribution des différents prédicteurs continus en un certain nombre de catégories d'effectifs sensiblement égaux. Pour les prédicteurs catégoriels, les catégories (classes) sont définies "naturellement".

2) Regroupement de catégories :

Ensuite il va examiner les prédicteurs afin de déterminer pour chacun, le couple de catégories (du prédicteur) les plus semblables (c'est-à-dire significativement moins différentes) par rapport à la variable dépendante ; pour les problèmes de classification (où la variable est également catégorielle), le programme va calculer un test du Chi^2 (*Chi-deux* de Pearson) , pour les problèmes de régression (où la variable dépendante est continue) le programme va calculer des tests F.

Si le test respectif, pour un couple donné de catégories du prédicteur, ne peut être considéré comme significatif ayant égard à une valeur alpha-de-fusion, l'algorithme va alors regrouper les catégories correspondantes du prédicteur et répéter ce processus (c'est-à-dire, rechercher le couple suivant de catégories, qui à présent peuvent être des catégories précédemment fusionnées). Si le couple respectif de catégories du prédicteur est statistiquement significatif (inférieur à la valeur correspondante du alpha-de-fusion), le programme va alors calculer une valeur- p ajustée de Bonferroni pour l'ensemble des catégories du prédicteur respectif.

3) Sélection d'une variable de division :

Enfin l'algorithme va choisir pour la division, la variable prédictive qui possède la plus faible valeur- p ajustée, c'est-à-dire la variable prédictive qui permet de produire la division la plus significative ; si la plus petite valeur- p ajustée (Bonferroni) des prédicteurs est supérieure à une certaine valeur alpha-de-division, le processus de division prend fin, et le nœud respectif est un nœud terminal.

4) Etape de l'arrêt de l'algorithme :

Le processus d'arrêt de l'algorithme se fait en accord avec un certain nombre de règles prédéfinies. Une seule de ces règles suffit.

- Si tous les nœuds sont purs, autrement dit si ils contiennent des modalités identiques de la variable dépendante.
- Si l'arbre a atteint la limite de croissance spécifiée par l'utilisateur.
- Si la taille d'un nœud est inférieure à la taille minimale spécifiée par l'utilisateur.

La p-value ajustée est un élément important de l'algorithme, il sert tant à fusionner des modalités qu'à les séparer. Cette p-value est calculée de deux manières différentes selon que la variable dépendante est une variable catégorielle ou une variable continue :

Dans le cas où c'est une variable continue, la p-value est obtenue à partir d'un test d'ANOVA. En effet il s'agira de déterminer si les moyennes dans les différentes catégories testées sont égales ou non. La statistique calculée est le rapport entre la moyenne des carrés intra-groupes et la moyenne des carrés intergroupes. Elle suit une Fisher ($g - 1, N - g$) où g est le nombre de groupes et N le nombre d'observations.

$$MC_{inter} = \frac{1}{g-1} \sum n_i \sum (y_i - y_{..})^2 \quad MC_{inter} = \frac{1}{N-g} \sum \sum (y_{ij} - y_i)^2$$

Avec : - y_{ij} est la j^{eme} observation du groupe i
 - n_i le nombre d'observations du groupe i

$$F = MC_{inter} / MC_{intra}$$

Dans le cas où c'est une variable nominale, la p-value est calculée à partir de la statistique du KHI-DEUX obtenue lors du test d'indépendance entre la variable dépendante et le prédicateur. Si la variable dépendante a I modalités et le prédicateur J modalités le KHI-DEUX de Pearson se calcule comme suit :

$$\sum_I \sum_J \frac{(n_{..}n_{ij} + n_i.n_j)^2}{n_{..}n_i.n_j}$$

Où n_{ij} est le nombre d'observations de niveau i pour la variable dépendante et de niveau j pour le prédicateur

III.2) Segmentation des variables :

✓ Le plafond Général :

L'algorithme CHAID nous segmente la variable Plafond Général par rapport à la variable charge en six classes qui se différencient par leur variance intra et inter-groupe. La figure ci-dessous nous donne l'arbre de segmentation avec une p-value très significatives.

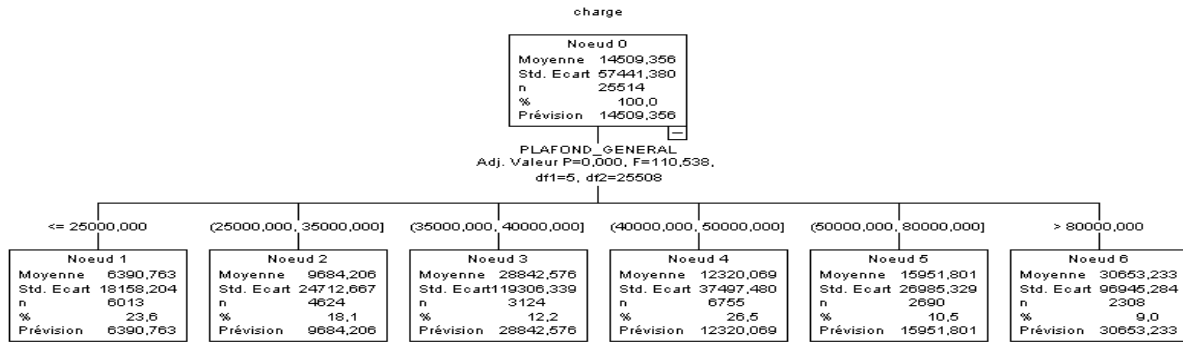


Figure 7-Arbre de codification du plafond général

Ainsi la codification de la variable plafond Général sera basée sur le résultat de l’algorithme CHAID.

✓ Forfait Naissance Simple :

La variable Forfait Naissance Simple a été segmentée en cinq classes avec une p-value très significative.

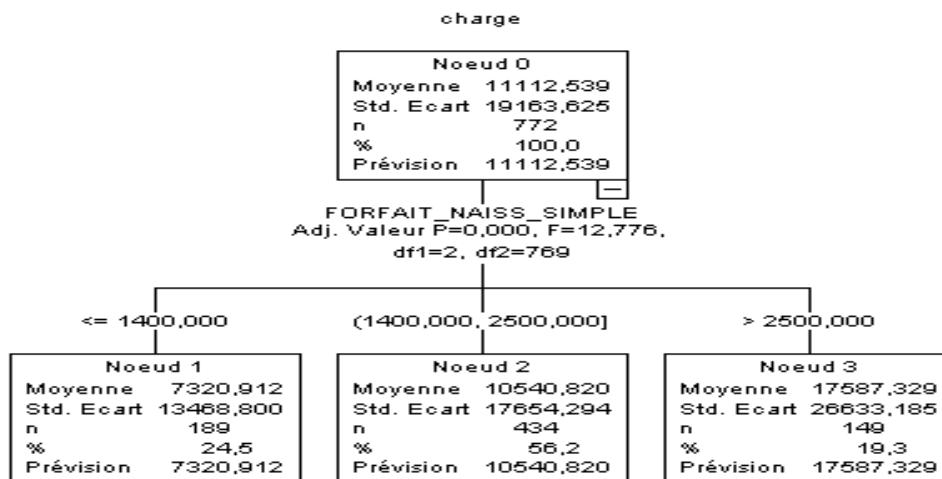


Figure 8-Arbre de codification Naissance simple

✓ Les autres variables Tarifaires:

Le tableau suivant résume la codification de nos variables :

Effectif du groupe	Code	1	2	3	4	5	6	7	8	9
	Segment	<=9] 9-14]] 14-31]] 31-49]] 49-70]] 70-107]] 107-190]] 190-401]	>401

Plafond général	Code	1	2	3	4	5	6
	Segment	<=25000] 25000- 35000]] 35000- 40000]] 40000 - 50000]] 50000 - 80000]	>80000
forfait naissance simple	Code	1		2		3	
	Segment	<=1400] 1400-2500]		>2500	
plafond prothèse et parodontie	Code	1		2		3	
	Segment	<=2800] 2800-3500]		>3500	
monture	Code	1		2		3	
	Segment	<=500] 500-750]		>750	

Tableau 20: Codification de nos variables tarifaires

Une fois la segmentation de nos variables tarifaires achevées, nous allons effectuer une statistique descriptive sur ces dernières.

IV. Statistiques descriptives sur le portefeuille :

Une analyse statistique descriptive du portefeuille est toujours utile pour permettre de comprendre et d'interpréter les résultats que pourrons nous donner les calculs. Le tableau suivant présente la répartition des effectifs assurés en 2013 et 2014 par type de bénéficiaire.

✓ Répartition des effectifs par type de bénéficiaire :

Type de bénéficiaire	2013		2014		Evolution
	Effectif	Poids	Effectif	poids	
Assurés	126 817	47.35%	133 912	48.00%	5,9%
Conjoints	52 066	19.44%	54 288	19.46%	4,3%
Enfants	88 928	33.21%	90 811	32.55%	2,1%

Tableau 21-Répartition des effectifs par type de bénéficiaire

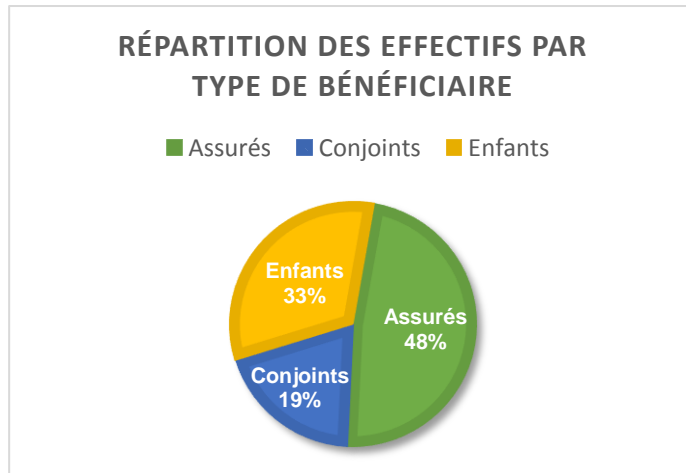


Figure 9-Répartition des effectifs par type de bénéficiaires

Les assurés principaux représentent la majeure partie des effectifs assurés tant en 2013 qu'en 2014 (presque 48 % durant ces deux années) avec une évolution de 5.9% de ceux-ci. Les conjoints représentent presque 20 % des effectifs avec une évolution de 4.3%, alors que les enfants représentent 33% des effectifs en 2013 et 32% en 2014 avec une évolution de 2.1%.

✓ Répartition des bénéficiaires suivant l'âge :

Les statistiques contenues dans le tableau suivant représentent la répartition des adhérents suivant l'âge moyen des différents types de bénéficiaires :

Type de bénéficiaire	Age Moyen	
	2013	2014
Assurés	41	40
Conjointes	40	39
Enfants	22	22

Tableau 22- Age moyen par type de bénéficiaire

L'âge moyen des assurés principaux s'élève à 40 ans en 2014, les conjoints ont en moyenne le même âge que celui des assurés (39 ans en 2014). Enfin, l'âge moyen des enfants est de 22 ans.

✓ Répartition des effectifs par classe du taux de remboursement :

Taux de remboursement	effectif	
	Somme	Pourcentage
<80	1 915 632.00	43.06%
80	1 126 525.00	25.32%
] 80,85]	828 583.00	18.62%
] 85,90]	337 202.00	7.58%
>90	241 019.00	5.42%

Tableau 23-Répartition des effectifs par classe du taux de remboursement

Nous constatons d'après le tableau ci-dessus que la partie majoritaire du portefeuille (43.06%) ont un taux de remboursement inférieur à 80%, suivi par une classe représentant 25.32% du portefeuille ayant ainsi un taux de remboursement égal à 80.

✓ Répartition des effectifs par classe du plafond général :

Plafond Général	effectif	
	Somme	Pourcentage
1	624 138.00	14.03%
2	467 060.00	10.50%
3	758 485.00	17.05%
4	1 086 366.00	24.42%
5	344 444.00	7.74%
6	1 168 468.00	26.26%

Tableau 24-Répartition des effectifs par classe du plafond général

Nous pouvons ainsi constater que la partie majoritaire du portefeuille (26.26%) ont un plafond général supérieur à 80000.

✓ Répartition des frais engagés par type de bénéficiaire :

Type de bénéficiaire	Frais engagés				Evolution
	Exercice 2013		Exercice 2014		
	Somme	pourcentage	Somme	pourcentage	
Assurés	113 455 527	58.60%	104 793 946	59.25%	-7,6%
Conjoints	53 695 469	27.73%	47 553 692	26.89%	-11,4%
Enfants	26 452 326	13.66%	24 524 650	13.87%	-7,3%

Tableau 25-Répartition des frais engagés par type de bénéficiaire

Nous pouvons constater que les assurés représentent la plus grande part de consommation, ce qui paraît logique vu qu'ils représentent la partie majeure des effectifs assurés, quoique la consommation des différents bénéficiaires a chuté de 7.6% ,11.4% et 7.3% par rapport à la consommation des assurés, conjoints et enfants respectivement en 2013.

✓ Répartition des frais engagés par type de prestation :

Type de prestation	Frais engagés				Evolution
	Exercice 2013		Exercice 2014		
	Somme	pourcentage	Somme	pourcentage	
Analyse (Biologie)	12 621 443	6.52%	12 442 822	7.03%	-1,4%
Dentaire	13 277 217	6.86%	10 871 552	6.15%	-18,1%
Hospitalisation	39 001 283	20.14%	34 840 565	19.70%	-10,7%
MONTURE	6 154 663	3.18%	5 524 106	3.12%	-10,2%
Maternité	6 058 132	3.13%	5 758 865	3.26%	-4,9%
NAISSANCE	4 088 727	2.11%	4 758 930	2.69%	16,4%
O.D.F	817 378	0.42%	821 885	0.46%	0,6%
Optique	10 554 009	5.45%	9 490 815	5.37%	-10,1%
PARODONTIE	117 220	0.06%	128 132	0.07%	9,3%
PROTHESE DENTAIRE	18 684 500	9.65%	18 006 746	10.18%	-3,6%
Pharmacie	40 214 562	20.77%	33 665 220	19.03%	-16,3%
Radiologie	9 506 358	4.91%	8 833 857	4.99%	-7,1%
Soins Ambulatoires	32 030 852	16.54%	30 976 583	17.51%	-3,3%
appareils orthopédiques	447 339	0.23%	717 377	0.41%	60,4%

Tableau 26-Répartition des Frais engagés par prestation

Nous pouvons d'ores et déjà constater que les prestations « Hospitalisation » et « Pharmacie » représentent la grande part de consommation avec un pourcentage de 19.77% et 19.03 % respectivement en 2014, suivi par la consommation en « soins ambulatoires » représentant 17.51% de la consommation totale, quoique la consommation relatives aux différentes prestations a diminué en 2014 par rapport à celle de 2013.

V. Le Cadre théorique du Modèle linéaire généralisé :

Dans cette partie nous présenteront le modèle linéaire généralisé ainsi que son application pour réaliser une tarification.

V.1) Rappel du modèle linéaire simple :

Le modèle linéaire simple consiste à établir une relation du type linéaire entre une variable à expliquer (aléatoire) et des variables explicatives (déterministes). L'équation du modèle est donc de la forme :

$$y = \beta_0 + \sum_{i=1}^p \beta_i X_i + \varepsilon$$

Où :

- Y est la variable à expliquer ;
- ε est une variable aléatoire de loi normale de moyenne nulle et de variance σ^2 qui représente l'écart entre le modèle et les observations ;
- $(X_1 X_2 \dots X_p)$ sont les variables explicatives déterministes ;
- $(\beta_0 \beta_1 \dots \beta_p)$ sont les paramètres du modèle à estimer ;

Ce modèle suppose donc que la variable à expliquer suit une loi normale de moyenne

$(\beta_0 + \sum_{i=1}^p \beta_i X_i)$ et de variance σ^2 .

La condition sur ε est difficile à satisfaire et demande des tests de normalité. Un des avantages du Modèle Linéaire Généralisé (GLM) est de supprimer cette contrainte.

V.2) Distribution de la famille exponentielle naturelle :

Les distributions de la famille exponentielle naturelle sont indispensables pour la mise en place d'un MLG. C'est pourquoi, nous allons présenter ici les principales distributions qui sont généralement utilisées.

✓ Forme générale :

Soit Y une variable aléatoire et y une observation de Y. La loi de probabilité de Y appartient à la famille exponentielle naturelle si et seulement si elle peut se mettre sous la forme :

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

Où :

- a est une fonction non nulle définie sur R ;
- b est une fonction non nulle définie sur R , deux fois dérivable ;
- c une fonction définie sur R^2 ;
- θ paramètre canonique ou paramètre de la moyenne ;
- φ est le paramètre de dispersion.

L'espérance de Y s'écrit alors $\mu = E(Y) = b'(\theta)$

On introduit la fonction de lien canonique g_c telle que : $\theta = g_c(\mu)$

La variance de Y s'écrit : $Var(Y) = b''(\theta) * a(\varphi)$

La fonction variance se définit par : $b''(\theta) = V(\mu)$

✓ Loi normale :

Soit Y une variable aléatoire suivant une loi normale d'espérance μ et de variance σ^2 . Y est à valeurs réelles. Sa fonction de densité est :

$$f_{\mu,\sigma}(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\}$$

Qui peut être mis sous la forme :

$$f_{\mu,\sigma}(y) = \exp\left\{\frac{y\mu - \frac{\mu^2}{2}}{\sigma^2} - \frac{\frac{y^2}{2} + \ln(2\pi\sigma^2)}{2}\right\}$$

Ainsi la loi gaussienne appartient à la famille exponentielle naturelle avec :

- $\theta = \mu$
- $\varphi = \sigma^2$
- $a(\varphi) = \varphi$
- $b(\theta) = \theta^2/2$
- $c(y; \varphi) = -\frac{1}{2}\left[\frac{y^2}{\varphi} + \ln(2\pi\varphi)\right]$

✓ Loi de Poisson :

Soit Y une variable aléatoire suivant une loi de Poisson de paramètre λ . Y Est à valeurs discrètes. Sa fonction de densité est de la forme :

$$P(Y = y) = \exp(-\lambda) \frac{\lambda^y}{y!}$$

Qui peut être mis sous la forme :

$$P(Y = y) = \exp(y \ln(\lambda) - \lambda - \ln(y!))$$

Ainsi la loi de Poisson appartient à la famille exponentielle naturelle avec :

- $\theta = \ln(\lambda)$
- $\varphi = 1$
- $a(\varphi) = 1$
- $b(\theta) = \exp(\theta)$
- $c(y; \varphi) = -\ln(y!)$

✓ Loi binomiale négative :

Soit Y une variable aléatoire suivant une loi binomiale négative de paramètres r et p (r étant un entier strictement positif, p un réel compris entre 0 et 1). Y est à valeurs entières. Sa fonction de densité est :

$$P(Y = y) = \exp(y \ln(\lambda) - \lambda - \ln(y!))$$

Qui peut être mise sous la forme :

$$P(Y = y) = \exp \left(y \ln (1 - p) + r \ln p + \ln \frac{y^{r-1}}{\Gamma(r)} \right)$$

Où

$$\Gamma(x) = \int_0^{\infty} e^{-u} u^{x-1} du$$

Ainsi la loi binomiale négative appartient à la famille exponentielle naturelle avec :

- $\theta = \ln (1 - p)$
- $\varphi = 1$
- $a(\varphi) = 1$
- $b(\theta) = -r \ln(1 - e^{\theta})$
- $c(y; \varphi) = -\ln \left(\frac{y^{r-1}}{\Gamma(r)} \right)$

✓ Loi Gamma :

Soit Y une variable aléatoire suivant une loi Gamma de paramètre r et α (tous deux strictement positifs). La densité s'écrit :

$$f_{r,\alpha}(y) = \frac{\alpha^r}{\Gamma(r)} y^{r-1} \exp(-\alpha y)$$

Où

$$\Gamma(x) = \int_0^{\infty} e^{-u} u^{x-1} du$$

La fonction de densité peut également être mise sous la forme :

$$f_{r,\alpha}(y) = \exp(-r \ln \alpha - \alpha y + (r - 1) \ln(y) - \ln \Gamma(\alpha))$$

Ainsi la loi Gamma appartient à la famille exponentielle naturelle avec :

- $\theta = -\alpha$
- $\varphi = 1$
- $a(\varphi) = 1$
- $b(\theta) = -r \ln(-\theta)$
- $c(y; \varphi) = (r - 1) \ln y - \ln \Gamma(\alpha)$

✓ Cas particulier de la loi log-normale :

Soit Y une variable aléatoire suivant une loi log-normale de paramètres μ et σ^2 .

Y est à valeurs réelles. Sa fonction de densité est :

$$f_{\mu,\sigma}(y) = \frac{1}{y\sigma\sqrt{2\pi}} \exp \left\{ -\frac{(\ln y - \mu)^2}{2\sigma^2} \right\}$$

Cette loi n'est pas une exponentielle naturelle. On ne peut donc pas appliquer un MLG sur la Variable Y directement. L'astuce consiste à poser $X = \ln(Y)$. X suit alors une loi normale de paramètres μ et σ^2 .

V.3) Hypothèse du modèle linéaire généralisé :

1) Hypothèses :

Le modèle linéaire généralisé part du même principe que celui du modèle linéaire simple. La différence est qu'au lieu de modéliser la variable à expliquer directement, c'est une fonction de l'espérance de cette variable (appelée fonction lien) qui est modélisée. Une variable aléatoire Y relève du modèle linéaire généralisé si la loi de Y sachant $\{X_1, X_2, \dots, X_p\}$ est telle que :

- Il existe une fonction lien g strictement monotone de \mathbb{R} dans \mathbb{R} et des coefficients $(\beta_0, \beta_1, \dots, \beta_p)$ tels que :

$$g(E[Y]) = \beta_0 + \sum_{i=1}^p \beta_i X_i$$

- La loi de probabilité de Y doit appartenir à la famille exponentielle naturelle
- Les paramètres $(\beta_0, \beta_1, \dots, \beta_p)$ sont les coefficients de régression et la quantité $g(E[Y])$ est le prédicteur linéaire.

2) Fonctions de lien :

Les fonctions de lien classiques sont les suivantes :

Fonction identité $g : z \rightarrow z$

Fonction logarithmique $g : z \rightarrow \ln(z)$

Fonction inverse $g : z \rightarrow 1/z$

En effet, si la fonction de lien est l'identité, on a :

$$E[Y] = \beta_0 + \sum_{i=1}^p \beta_i X_i$$

Ou encore :

$$E[Y] = \sum_{i=1}^p \alpha_i$$

Avec :

- $\alpha_0 = \beta_0$
- $\alpha_i = \beta_i X_i$ pour tout i entre 1 et p

Il s'agit bien d'un modèle Additif.

Si la fonction de lien est le logarithme, on a :

$$E[Y] = \exp \left(\beta_0 + \sum_{i=1}^p \beta_i X_i \right)$$

Ou encore

$$E[Y] = \prod_{i=1}^p \alpha_i$$

Avec

- $\alpha_0 = \exp(\beta_0)$
- $\alpha_i = \exp(\beta_i X_i)$ pour tout i entre 1 et p

Il s'agit bien d'un modèle multiplicatif.

V.4) Estimation des paramètres :

L'estimation des paramètres du MLG se fait par maximum de vraisemblance. Prenons l'exemple d'une loi exponentielle de paramètre θ et de densité f . Soit Y une variable aléatoire suivant cette loi et (y_1, y_2, \dots, y_n) p observations de cette variable.

Rappelons que notre modèle est de la forme :

$$g(\mu) = \beta_0 + \sum_{i=1}^p \beta_i X_i$$

Avec

$$\mu = E[Y]$$

Nous avons de plus :

$$\theta = g_c(\mu)$$

Si on note la fonction de vraisemblance :

$$L(y_1, y_2, \dots, y_n, \beta_0, \beta_1, \dots, \beta_p) = \prod_{i=1}^n f(y_i, \theta)$$

L'équation à résoudre sera :

$$\frac{\partial L(y_1, y_2, \dots, y_n, \beta_0, \beta_1, \dots, \beta_p)}{\partial \alpha} = \sum_{i=1}^p \frac{\partial \ln f(y_i, \beta_0, \beta_1, \dots, \beta_p)}{\partial \alpha} = 0$$

On obtient ainsi $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p)$ les estimations des paramètres du modèle. Ces estimations nous seront données par la suite par le logiciel SAS et la procédure GENMOD

V.5) Significativité des variables :

La significativité des coefficients associés aux variables explicatives peut être testée à l'aide du test de Wald.

Soit le test suivant :

$$H_0: \beta = 0 \text{ Contre } H_1: \beta \neq 0$$

Ainsi la statistique de Wald s'écrit :

$$W = \frac{\hat{\beta}_j}{\hat{\sigma}(\hat{\beta}_j)}$$

Sous H_0 , la statistique du test suit approximativement une loi Normale $N(0, 1)$. La test de Wald peut aussi être définie ainsi :

$$Z = \left(\frac{\hat{\beta}_j}{\hat{\sigma}(\hat{\beta}_j)} \right)^2$$

Sous H_0 , cette statistique suit asymptotiquement une loi de khi-deux à un degré de liberté.

✓ Règle de décision

Au niveau de signification α : on rejette H_0 si $|W| > \phi(1 - \alpha)$ avec : ϕ la fonction réciproque de la loi normale $N(0,1)$.

V.6) Qualité du modèle :

Il existe plusieurs critères pour juger la qualité d'ajustement du modèle :

- *La déviance* : permet de comparer le modèle estimé avec le modèle dit saturé (Modèle où il y'a autant d'observations que de paramètres). Ce modèle saturé estime donc précisément les données. La déviance est définie à partir de la log-vraisemblance de ces deux modèles :

$$D = -2 (L - L_{sat})$$

La déviance D suit asymptotiquement une loi de khi 2 à $n-p-1$ degré de liberté.

Un test de rejet ou d'acceptation du modèle basé sur la déviance peut être utilisé :

Si la déviance est supérieure au quantile d'une loi de khi 2 à $n-p-1$ degrés de liberté d'ordre $1 - \alpha$, le modèle est jugé de mauvaise qualité.

- *La Statistique de Pearson* : permet de comparer les valeurs observées Y_i à leur prévision par le modèle. Cette statistique est définie par :

$$X^2_{pearson} = \sum_{i=1}^n \frac{(Y_i - \hat{\mu}_i)^2}{\widehat{var}(\hat{\mu}_i)}$$

La statistique de Pearson X^2 suit asymptotiquement une loi de khi 2 à $n-p-1$ degrés de liberté. Un test de rejet ou d'acceptation du modèle basé sur cette statistique peut être utilisé :

Si X^2 est supérieure au quantile d'une loi de khi deux à $n-p-1$ degrés de liberté d'ordre $1 - \alpha$, le modèle est jugé de mauvaise qualité.

V.7) Sélection des variables :

Une sélection des variables doit être effectuée afin de trouver le meilleur modèle. Pour classer les modèles, trois critères peuvent être utilisés :

- Le critère « LL » égal à $-2\log(L)$ avec L est le maximum de vraisemblance.
- Le critère « AIC » égal à $-2\ln(L)+2p$ ou p est le nombre de variables explicatives.
- Le critère « BIC » égal à $-2\ln(L)+p\ln(n)$ ou n est le nombre d'observations.

Plus ces critères sont petits, plus le modèle est considéré comme bon.

Les méthodes les plus utilisées de sélection sont les suivantes :

La méthode ascendante (BACKWARD) : il s'agit de démarrer avec le modèle complet (c'est-à-dire toutes les variables ayant un effet significatif sur le risque) puis de retirer la variable la moins significative, autrement dit celle dont l'élimination entraîne la plus faible augmentation de la déviance.

Dans la méthode ascendante (Forward) : il s'agit de rechercher la variable la plus significative au sens du R2 pour un modèle linéaire classique ou au sens de la déviance pour un modèle linéaire généralisé. Partant de ce modèle à un facteur, nous cherchons ensuite la variable qui, associée à la première, explique le mieux la sinistralité et ainsi de suite.

V.8) Validation du modèle :

L'analyse des résidus permet une analyse plus poussée que les statistiques vues dans le paragraphe précédent. Elle permet en effet de comprendre d'où proviennent les éventuels écarts entre les valeurs prédites $\hat{\mu}_i$ et les données en détectant les observations particulières.

Deux types de résidus sont classiquement utilisés pour les modèles linéaires généralisés : les résidus de Pearson et les résidus de la déviance.

✓ Résidus de Pearson :

$$r_i^p = \frac{Y_i - \hat{\mu}_i}{\sqrt{\widehat{var}(\hat{\mu}_i)}}$$

✓ Résidus de la déviance :

$$r_i^d = \text{signe}(Y_i - \hat{\mu}_i)\sqrt{d_i}$$

d_i Est la composante de la déviance induite par la $i^{\text{ème}}$ observation, avec $D = \sum_1^n d_i$ la déviance du modèle. C'est cette dernière approche que nous utilisons dans les calculs.

La validité du modèle est jugée bonne si les résidus observés se situent autour de l'axe des abscisses et avec une variance constante selon i , autrement dit si le nuage de points est de forme cylindrique autour de l'axe des abscisses.

VI. Application de la tarification à l'aide d'un modèle linéaire généralisé :

Nous étudions dans cette partie la façon dont les modèles linéaires généralisés sont appliqués à notre étude. Ainsi, après avoir rappelé le cadre théorique des modèles linéaires généralisés, nous présenterons par la suite les modélisations qui sont retenues pour nos deux variables à savoir la fréquence et le cout moyen des sinistres.

Comme nous l'avons mentionné déjà le risque maladie comporte plusieurs types de prestations comme pharmacie, Radiologie, hospitalisation dentaire et optique etc.... Pour modéliser le risque maladie nous allons modéliser chaque prestation à part. Ainsi, les différentes prestations que nous tariferons sont :

- Analyse (Biologie) ;
- Pharmacie ;
- Parodontie et Prothèses dentaire (Assuré et Conjoint) ;
- soins Ambulatoires ;
- Optique ;
- Odf, Parodontie et Prothèses dentaire (enfant) ;
- Radiologie

- Naissance
- Maternité
- Monture
- Hospitalisation
- Dentaire

Pour chacun de ces groupes de prestations nous avons introduit les variables suivantes qui nous seront utiles dans notre modélisation :

- $\ln(\text{Charge/nombres de sinistres})$ qui est la variable à expliquer dans le cas d'un modèle linéaire généralisé supposant une loi log normale du coût moyen.
- $\ln(\text{exposition*effectif})$ qui est la variable offset dans le cas de la modélisation de la fréquence des sinistres avec une distribution poissonnienne ou binomiale négative avec une fonction de lien logarithmique.

Par la suite, nous détaillerons la démarche de la tarification pour la prestation « pharmacie » pour chaque type de bénéficiaires, et pour les autres prestations, nous établirons en annexe 4 les résultats des modèles choisis.

VI.1. Modélisation de la garantie « Pharmacie » :

1) Analyse descriptive :

Dans cette partie nous allons donner la fréquence et le coût moyen des sinistres pour chaque modalité de chaque variable qui pourrait être significative pour notre modèle.

✓ Lien :

Le tableau ci-dessous représente le coût moyen ainsi que la fréquence des sinistres de chaque modalité et leur poids, on remarque donc que les assurés sont les plus risqués en termes de coût moyen des sinistres que les autres modalités avec un poids de 49.24%.

<i>Lien</i>	<i>Coût Moyen</i>	<i>Fréquence</i>	<i>Poids</i>
Assuré	370.48	1.29	48.00%
Conjoint	317.17	1.00	19.46%
Enfant	190.86	0.56	32.55%

Tableau 27- Coût moyen et fréquence par type de bénéficiaire

✓ Taux de remboursement :

On remarque donc que la modalité 5 référant au taux de remboursement supérieur à 90 est la plus risquée en termes du coût moyen des sinistres que les autres modalités et qui constitue une partie minoritaire avec un pourcentage de 5.42% de la population étudiée.

<i>Taux de remboursement</i>	<i>Coût Moyen</i>	<i>Fréquence</i>	<i>Poids</i>
<80	295.50	0.83	43.06%
80	292.26	0.92	25.32%
] 80,85]	299.93	1.02	18.62%
] 85,90]	286.93	1.62	7.58%
>90	535.22	1.10	5.42%

Tableau 28-Cout moyen et fréquence par taux de remboursement

✓ Plafond général:

Nous pouvons constater que la classe ayant un plafond général supérieur à 80 000 est celle qui représente le plus de risque en termes du cout moyen quoiqu'elle représente ainsi la partie majoritaire dans le portefeuille étudié.

<i>Plafond Général</i>	<i>Coût Moyen</i>	<i>Fréquence</i>	<i>Poids</i>
<=25 000	300.35	0.92	14.03%
] 25 000,35 000]	261.51	0.89	10.50%
] 35 000 ,40 000]	285.30	0.90	17.05%
] 40 000,50 000]	309.19	1.32	24.42%
] 50 000,80 000]	271.99	0.95	7.74%
>80000	495.54	0.80	26.26%

Tableau 29-Cout moyen et fréquence par plafond général

✓ Effectif du groupe :

D'après le tableau ci-dessous, nous pouvons ainsi constater que la modalité qui ne constitue qu'une partie très minoritaire du portefeuille est celle qui représente le maximum de risque en termes du cout moyen. Par ailleurs, la partie majoritaire du portefeuille (66.99%) est celle dont l'effectif est supérieur à 401 et a un coût moyen de 286.77 Dhs.

<i>Effectif du groupe</i>	<i>Coût Moyen</i>	<i>Fréquence</i>	<i>Poids</i>
>=9	379.78	1.94	0.35%
]9-14]	337.16	1.02	0.58%
]14-31]	296.27	0.85	2.27%
]31-49]	282.54	0.78	2.34%
]49-70]	301.05	0.83	3.44%

]70-107]	253.63	0.82	4.81%
]107-190]	277.12	0.76	8.33%
]190-401]	285.38	0.68	10.88%
>401	286.77	0.46	66.99%

Tableau 30-Cout moyen et fréquence par effectif de groupe

Par la suite, nous retenons les variables explicatives suivantes :

- le taux de remboursement ;
- Le plafond général ;
- L'effectif du groupe (pour modéliser ainsi la prime des assurés principaux)
- Les sous plafonds

Après Segmentation, des nouvelles variables ordinales seront créées prenant comme valeur les modalités des variables tarifaires. Ces variables seront nommées comme suit :

- le taux de remboursement : tranche_tr
- Le plafond général : tranche_pg
- L'effectif du groupe (pour modéliser ainsi la prime des assurés principaux) : tranche_na
- Les sous plafonds : Le plafond monture : tranche_mt ; Le plafond Odf, parodontie et prothèse dentaire : tranche_ppd ; Le Forfait Naissance simple : forfait_ns

2) Modélisation du Cout Moyen des « Assurés » :

a) Ajustement et choix du modèle :

Plusieurs techniques sont utilisées afin de choisir la loi la plus adaptée aux données, celle qui s'y ajuste le mieux. Nous choisirons une technique graphique, Q-Q plot (quantile-to-quantile plot). Cette technique compare entre les quantiles empiriques et les quantiles théoriques de la loi qui ajuste les données. Sous SAS, nous appliquerons les tests des Q-Q plot avec les lois candidates et choisirons celle où les points seront proches de la bissectrice. Les deux modèles les plus classiques permettant de modéliser le coût moyen des sinistres sont le modèle Gamma et le modèle log-normal. Les deux graphiques suivants montrent l'ajustement des deux lois avec la distribution réelle des charges.

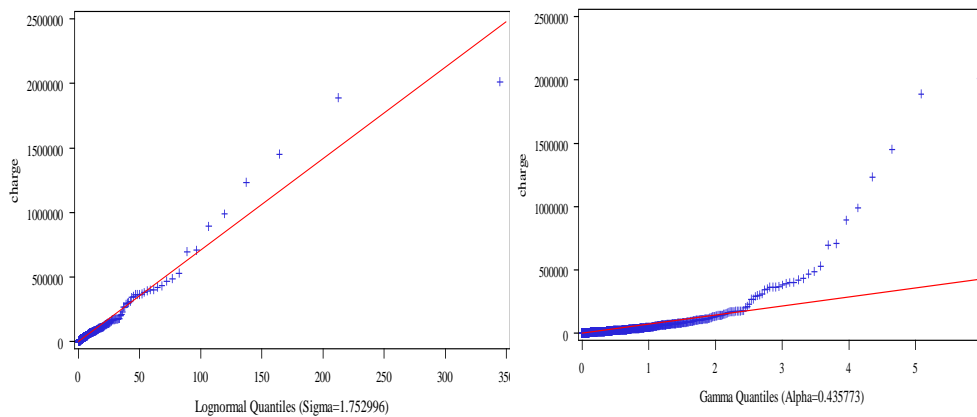


Figure 10-QQ-plot des lois Gamma et log normale

D'après la figure ci-dessus, nous constatons que la loi log normale s’ajuste mieux que la loi gamma. Donc nous avons retenu la loi log normale pour la consommation de la prestation « Pharmacie » des assurés. De la même manière nous avons procédé au choix de la loi lognormale pour les autres types de bénéficiaires (voir annexe 4).

Ainsi, en se basant sur les résultats de la déviance, il paraît ainsi que la loi lognormale est celle qui s’ajuste le mieux aux données puisque la valeur des deux statistiques est proche de leur degré de liberté car l’espérance d’une loi Khi-Deux égale à son degré de liberté hors les deux statistiques suivent asymptotiquement une loi Khi-Deux.

Loi		Gamma		Log Normale		
Critère	DF	Value	$\frac{Valeur}{DF}$	DF	Value	$\frac{Valeur}{DF}$
Déviance	1437	0.0000	0.0000	1437	1455.0000	1.0125
Statistique de pearson	1437	4.364986E19	3.0375685E16	1437	1455.0000	1.0125

Tableau 31- Statistique de Déviance du modèle par loi

b) Sélection des variables :
 ✓ Test de type 3 :

La procédure GENMOD de SAS fournit les déviances de vraisemblance (type 1 et type 3) fréquemment utilisées pour tester l’apport des facteurs. Ainsi, nous nous sommes basés sur le test de type 3 pour tester la significativité des variables explicatives. Ainsi le degré de significativité de chaque variable est présenté dans le tableau ci-dessous :

Statistiques LR pour Analyse de Type 3			
Source	DDL	Khi-2	Pr > Khi-2
tranche_tr	4	12.21	0.0135
tranche_na	8	15.60	0.0485
tranche_pg	5	26.88	<.0001

Tableau 32-Statistiques type 3 des variables du modèle ajusté au coût par une loi lognormale

Dans ce cas, le Khi deux indiqué dans la sortie de SAS ci-dessus correspond à la différence des déviations des modèles avec et sans la variable considérée sur la ligne. La dernière colonne indique la p-value associée au test du rapport des vraisemblances. Pour un niveau de confiance de 95%, par exemple, on accepte les variables qui donnent une p-value inférieure à 0,05. Concrètement, la différence des déviations des modèles avec la constante, la tranche_tr, la tranche_na et la tranche_pg d’une part, et la constante ,la tranche_tr et la tranche_na d’autre part vaut 26.88. Ceci montre l’apport de la variable plafond général indépendamment de l’ordre dans lequel elle apparaît. Ici, le plafond général semble plus significatif que les autres variables.

c) Résultats du modèle :

Sur la base de la procédure GENMOD sous sas, nous avons ainsi obtenu les résultats suivants pour une loi lognormale, cependant les valeurs estimées ainsi que leurs écarts types ne seront pas fournis pour des raisons de confidentialité.

Analyse des valeurs estimées du paramètre de vraisemblance maximum								
Paramètre	Code	DDL	Valeur estimée	Erreur Type	Intervalle de confiance Wald à 95%		Khi de Wald	Pr>Khi-2
Intercept		1	****	****	****	****	3780,1	<0.0001
tranche_tr	1	1	****	****	****	****	9,42	0,0012
tranche_tr	2	1	****	****	****	****	7,98	0,0047
tranche_tr	3	1	****	****	****	****	6,06	0,0057
tranche_tr	4	1	****	****	****	****	9,78	0,0017
tranche_tr	5	0	****	****	****	****		
tranche_na	1	1	****	****	****	****	4,87	0,0021
tranche_na	2	1	****	****	****	****	9,93	0,0082
tranche_na	3	1	****	****	****	****	6,91	0,0036
tranche_na	4	1	****	****	****	****	7,05	0,0207
tranche_na	5	1	****	****	****	****	9,45	0,0067
tranche_na	6	1	****	****	****	****	0,02	0,8925
tranche_na	7	1	****	****	****	****	9,22	0,0324
tranche_na	8	1	****	****	****	****	9,23	0,0358
tranche_na	9	0	****	****	****	****		
tranche_pg	1	1	****	****	****	****	7,62	0,0429
tranche_pg	2	1	****	****	****	****	20,783	<0.0001
tranche_pg	3	1	****	****	****	****	9,7523	0,0017
tranche_pg	4	1	****	****	****	****	8,0983	0,0044
tranche_pg	5	1	****	****	****	****	7,9997	0,0046
tranche_pg	6	0	****	****	****	****		

Scale		1	****	****	****	****		
-------	--	---	------	------	------	------	--	--

Tableau 33-les coefficients estimés du Modèle du cout des assurés pour la «Pharmacie »

Sur le tableau nous pouvons clairement lire que la sixième modalité de la variable tranche_na n'est pas significative, la p-value correspondante est égale à 0.8925. Mais avant de conclure, nous allons tout d'abord effectuer le test du rapport de vraisemblance, où nous testerons l'hypothèse d'égalité des coefficients des modalités de la variable tranche_na prises deux par deux pour voir s'il y a lieu d'effectuer des regroupements.

Par ailleurs, Les coefficients associés aux autres modalités semblent très significatifs d'après la statistique de Wald (pvalue < 0.05).

Le tableau ci-dessous, nous donne un résumé des tests effectués entre les différentes modalités :

Résultats de contraste				
Regroupement	DDL	Khi-2	Pr > Khi-2	Type
tranche_na 6&7	1	4.82	0.028	Wald
tranche_na 5&6	1	0.21	0.6442	Wald

Tableau 34-Test de contraste pour le regroupement des classes des modalités de la variable effectif

La cinquième et la sixième modalité de la variable tranche_na, qui correspondent respectivement aux groupes dont l'effectif est entre 49 et 70 et ceux ayant des effectifs entre 70 et 107 assurés, devraient être regroupées. Ainsi, les résultats du nouveau modèle sont les suivants :

Statistiques LR pour Analyse de Type 3			
Source	DDL	Khi-2	Pr > Khi-2
tranche_tr	4	12.21	0.0135
tranche_na	7	15.35	0.0317
tranche_pg	5	26.75	<.0001

Tableau 35- Statistiques type 3 des variables du modèle ajusté au coût par une loi lognormale après regroupement des modalités

Le tableau suivant nous présente les coefficients estimés ainsi que leurs significativités :

Analyse des valeurs estimées du paramètre de vraisemblance maximum								
Paramètre		DDL	Valeur estimée	Erreur Type	Intervalle de confiance Wald à 95%		Khi de Wald	Pr>Khi-2
Intercept		1	****	****	****	****	3904,97	<.0001
tranche_tr	2	1	****	****	****	****	3,92	0,0477
tranche_tr	3	1	****	****	****	****	7,41	0,0065
tranche_tr	4	1	****	****	****	****	3,93	0,0473
tranche_tr	5	0	****	****	****	****		
tranche_na	1	1	****	****	****	****	6,82	0,0090
tranche_na	2	1	****	****	****	****	5,47	0,0194
tranche_na	3	1	****	****	****	****	4,13	0,0421
tranche_na	4	1	****	****	****	****	3,85	0,0497

tranche_na	5	1	****	****	****	****	4,24	0,0395
tranche_na	7	1	****	****	****	****	7,52	0,0061
tranche_na	8	1	****	****	****	****	5,62	0,0178
tranche_na	9	0	****	****	****	****		
tranche_pg	2	1	****	****	****	****	10,53	0,0012
tranche_pg	3	1	****	****	****	****	9,68	0,0019
tranche_pg	4	1	****	****	****	****	8,58	0,0034
tranche_pg	5	1	****	****	****	****	7,78	0,0053
tranche_pg	6	0	****	****	****	****		
Scale		1	****	****	****	****		

Tableau 36-les nouveaux paramètres estimés après regroupement des modalités

Après la modification apportée aux deux modalités, les coefficients ont été ré-estimés. Nous remarquons que cette fois toutes les modalités sont significatives.

d) L'analyse des résidus :

On interprète les résultats sur un graphique en plaçant en ordonnée les observations des résidus de la déviance r_i et en abscisses les valeurs estimées $\hat{\mu}_i$. Pour que nous obtenions un modèle valide, il faut que les résidus soient assez proches de 0 et soient répartis de manière assez uniforme autour de l'axe des abscisses. Ainsi, dans notre exemple, Le graphique des résidus obtenu pour la prestation « pharmacie » modèle assuré est le suivant :

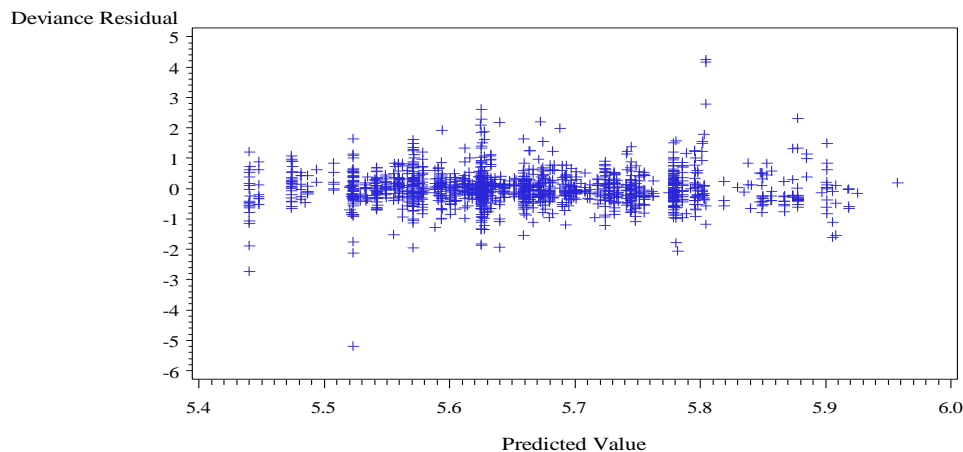


Figure 11- Représentation Graphique des résidus de la déviance

Les résidus sont correctement répartis sur l'axe des abscisses et sont d'une manière générale relativement proche de 0. On n'observe pas de valeurs étonnamment éloignée de 0 pour les résidus de la déviance, ce qui peut traduire une certaine justesse du modèle.

3) Modélisation de la Fréquence des « Assurés » :

a) Ajustement et choix du modèle :

Nous allons à présent, entamer la modélisation de la fréquence des sinistres. Classiquement en tarification, deux lois sont à notre disposition pour modéliser les fréquences : la loi de poisson et la loi binomiale négative. Nous devons donc choisir celle qui

s'ajuste le mieux aux données. Pour adopter le modèle adéquat, nous allons nous appuyer sur la déviance et la statistique de pearson, elle mesure la qualité d'ajustement du modèle testé.

Loi		Poisson		Binomiale Négative		
Critère	DF	Value	$\frac{Valeur}{DF}$	DF	Value	$\frac{Valeur}{DF}$
Déviance	1437	105263.7105	73.2524	1437	1562.3687	1.0872
Statistique de pearson	1437	126407.0144	87.9659	1437	1402.3059	0.9759

Tableau 37-Statistique de Déviance du modèle de fréquence par loi

Le tableau ci-dessous représente la déviance des modèles binomiale négative et poisson pour le poste « Pharmacie ».

Nous remarquons nettement que la loi binomiale négative est celle la mieux adaptée pour modéliser la fréquence du poste « Pharmacie ».

b) Sélection des variables :
 ✓ Test de type 3 :

Le degré de significativité de chaque variable est présenté dans le tableau ci-dessous :

Statistiques LR pour Analyse de Type 3			
Source	DDL	Khi-2	Pr > Khi-2
tranche_tr	4	10.27	0.036
tranche_na	8	205.14	<.0001
tranche_pg	5	15.50	0.0084

Tableau 38- Statistiques type 3 des variables du modèle ajusté de la fréquence par une loi Binomiale Négative

D'après ce tableau nous concluons que toutes les variables retenues sont significatives. Nous pouvons aussi remarquer l'apport très significatif de la variable « effectif » (tranche_na) dans le modèle. En effet, l'introduction de cette variable dans le modèle a fait diminuer la déviance de 205.14. Ainsi, cette dernière semble très significative par rapport aux autres variables, qui introduites dans le modèle n'ont fait diminué la déviance respectivement que de 10.07 et 15.5.

c) Résultats du modèle :

Sur la base de la procédure GENMOD sous sas, nous avons ainsi obtenu les résultats suivants pour une loi lognormale :

Analyse des valeurs estimées du paramètre de vraisemblance maximum								
Paramètre	Code	DDL	Valeur estimée	Erreur Type	Intervalle de confiance Wald à 95%		Khi de Wald	Pr>Khi-2
Intercept		1	****	****	****	****	21,93	<.0001
tranche_tr	1	1	****	****	****	****	0,58	0,4479
tranche_tr	2	1	****	****	****	****	11,15	0,0008
tranche_tr	3	1	****	****	****	****	3,72	0,0536
tranche_tr	4	1	****	****	****	****	15,28	0,0001

tranche_tr	5	0	****	****	****	****		
tranche_na	1	1	****	****	****	****	153,05	<.0001
tranche_na	2	1	****	****	****	****	48,13	<.0001
tranche_na	3	1	****	****	****	****	43,39	<.0001
tranche_na	4	1	****	****	****	****	26,02	<.0001
tranche_na	5	1	****	****	****	****	20,02	<.0001
tranche_na	6	1	****	****	****	****	17,53	<.0001
tranche_na	7	1	****	****	****	****	16,71	<.0001
tranche_na	8	1	****	****	****	****	5,87	0,0154
tranche_na	9	0	****	****	****	****		
tranche_pg	1	1	****	****	****	****	0,68	0,4095
tranche_pg	2	1	****	****	****	****	4,50	0,0338
tranche_pg	3	1	****	****	****	****	10,74	0,0010
tranche_pg	4	1	****	****	****	****	7,33	0,0068
tranche_pg	5	1	****	****	****	****	12,13	0,0001
tranche_pg	6	0	****	****	****	****		
Dispersion		1	****	****	****	****		

Tableau 39-paramètres estimés du modèle de fréquence pour « la Pharmacie »

Sur le tableau 39 nous pouvons lire que les premières modalités des variables tranche_pg et tranche_tr ne sont pas significatives, la p-value correspondante est égale à 0.4095 et 0.4479 respectivement. De la même manière qu’avec le cout moyen, nous allons effectuer le test du rapport de vraisemblance, où nous testerons l’hypothèse d’égalité des coefficients des modalités des variables tranche_pg et tranche_tr prises deux par deux pour voir s’il y a lieu d’effectuer des regroupements. Par ailleurs, Les coefficients associés aux autres modalités semblent très significatifs d’après la statistique de Wald (pvalue < à 0.05).

Le tableau ci-dessous, nous donne un résumé des tests effectués entre les différentes modalités :

Résultats de contraste				
Regroupement	DDL	Khi-2	Pr > Khi-2	Type
tranche_tr 1&2	1	1.39	0.238	Wald
Tranche_tr 1&3	1	5.91	0.015	Wald

Tableau 40-Test sur le regroupement des modalités de la variable taux de remboursement

Résultats de contraste				
Regroupement	DDL	Khi-2	Pr > Khi-2	Type
tranche_pg 1&2	1	4.17	0.041	Wald

Tableau 41-Test sur le regroupement des modalités de la variable plafond général

La première et la deuxième modalité de la variable tranche_tr, devraient être regroupées. Ainsi, les résultats du nouveau modèle sont les suivants :

Statistiques LR pour Analyse de Type 3			
Source	DDL	Khi-2	Pr > Khi-2
tranche_tr	3	9.09	0.028
tranche_na	8	205.14	<.0001
tranche_pg	4	16.42	0.0025

Tableau 42-Statistiques de type 3 après regroupement des modalités

Analyse des valeurs estimées du paramètre de vraisemblance maximum								
Paramètre	Code	DDL	Valeur estimée	Erreur Type	Intervalle de confiance Wald à 95%		Khi de Wald	Pr>Khi-2
Intercept		1	****	****	****	****	18,4752	<.0001
tranche_tr	2	1	****	****	****	****	4,89	0,0270
tranche_tr	3	1	****	****	****	****	4,53	0,0321
tranche_tr	4	1	****	****	****	****	15,11	0,0001
tranche_tr	5	0	****	****	****	****		
tranche_na	1	1	****	****	****	****	141,69	<.0001
tranche_na	2	1	****	****	****	****	41,41	<.0001
tranche_na	3	1	****	****	****	****	38,17	<.0001
tranche_na	4	1	****	****	****	****	22,51	<.0001
tranche_na	5	1	****	****	****	****	16,71	<.0001
tranche_na	7	1	****	****	****	****	19,85	<.0001
tranche_na	8	1	****	****	****	****	5,23	0,0222
tranche_na	9	0	****	****	****	****		
tranche_pg	2	1	****	****	****	****	6,62	0,0101
tranche_pg	3	1	****	****	****	****	9,87	0,0017
tranche_pg	4	1	****	****	****	****	7,02	0,0081
tranche_pg	5	1	****	****	****	****	4,16	0,0414
tranche_pg	6	0	****	****	****	****		
Dispersion		1	****	****	****	****		

Tableau 43-les nouveaux paramètres estimés après regroupement des modalités

Après la modification apportée aux deux modalités, les coefficients ont été ré-estimés. Nous remarquons que cette fois toutes les modalités sont significatives.

4) Calcul de la prime pure pour la loi log-normale :

Notre modélisation de la fréquence a porté sur un modèle multiplicatif avec une fonction lien logarithme, tandis que dans la modélisation du cout moyen, on a opté pour une fonction de lien identité puisqu'on a modélisé le log de la charge. Ainsi, la prime pure sera :

$$prime\ pure = \exp(\beta_0 + \sum_{i=1}^p \beta_i X_i) * \exp(\mu + \frac{\sigma^2}{2})$$

Avec : $\mu = \alpha_0 + \sum_{i=1}^p \alpha_i X_i$

Tels que : $(\beta_0 \beta_1 \dots \beta_p)$ sont les coefficients du modèle de la fréquence et $(\alpha_0 \alpha_1 \dots \alpha_p)$ sont les coefficients obtenus par le modèle du cout moyen ayant une fonction de lien identité.

Ainsi, une prime par prestation est calculée pour chaque type de bénéficiaires. Les résultats de la modélisation pour chaque type de prestation se trouvent en annexe 4.

✓ Présentation de l'agent de référence :

L'agent de référence est le tarif auquel nous référons lors du calcul du tarif des autres classes tarifaires, puisque les variables explicatives étant qualitatives (à K modalités), chaque modalité est introduite dans le modèle sous une forme binaire. Autrement dit, la variable correspondante à la modalité prend la valeur 1 si l'entreprise y appartient et zéro sinon. Les modalités ne seront pas introduites simultanément dans le modèle, une sera amputée et considérée comme classe de référence. Les modalités non introduites dans le modèle auront donc des coefficients nuls.

Notre classe de référence est la suivante :

Variable	Classe de référence
Plafond général	>80000
Taux de remboursement	>90
Effectif du groupe (na)	>401
Plafond odf	>3500
Plafond Monture	>350
Forfait naissance	>2500

Tableau 44- Caractéristiques de la classe de référence

Ainsi, la prime de référence aura l'expression suivante :

$$prime\ pure_{référence} = \exp(\beta_0) * \exp\left(\alpha_0 + \frac{\sigma^2}{2}\right)$$

Ainsi, une prime par prestation est calculée pour chaque type de bénéficiaire pour établir le tarif global du groupe.

VI.2. Méthode du calcul du tarif global :

Considérons une entreprise dont les caractéristiques de ses salariés sont les suivantes :

- L'effectif des salariés assurés égal à N_a
- L'effectif des bénéficiaires conjoints égal à : N_c
- L'effectif des bénéficiaires enfants est : N_e

Par ailleurs, selon les termes contractuels choisis par l'entreprise concernant le taux de remboursement, le plafond général ainsi que des autres plafonds, nous pouvons calculer la prime par prestation de chaque type de bénéficiaire.

Ainsi, la prime totale de chaque type de bénéficiaire sera :

$$prime\ totale_{type\ de\ bénéficiaire} = \sum_{prestation} prime_{prestation}$$

Par la suite, la prime qui sera payée par le groupe sera :

$$Prime = N_a * Prime\ totale_A + N_c * Prime\ totale_C + N_E * Prime\ totale_E$$

VII. Outil de tarification :

Après avoir calculé les différentes primes pures pour les différentes classes, nous avons été amenés à créer un outil de tarification clair et lisible par tous les agents et courtiers offrant les produits Saham Assurance. Afin de remplir cette tâche nous avons conçu une application VBA qui a comme input les caractéristiques générales du groupe et les caractéristiques propres à chaque prestation à savoir le taux de remboursement et les sous plafonds et comme output les différents tarifs associés à ces inputs.

L'application se présente comme suit, Dans un premier temps en ouvrant la feuille Excel une feuille s'ouvre avec un bouton « démarrer » qui apparaît automatiquement. La prise d'écran suivante représente ce bouton :

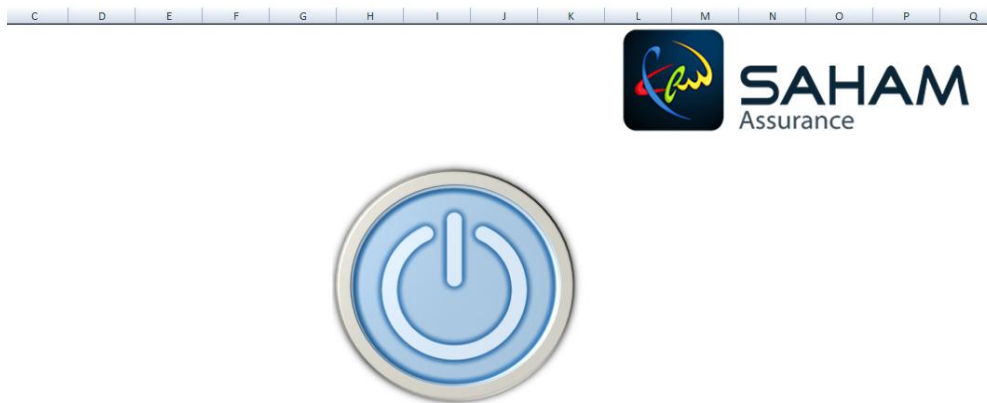


Figure 12: Bouton Démarrage d'application

En appuyant sur le bouton « démarrer », un userform multipage apparaît :

✓ La première page contient des labels portant les noms des variables tarifaires. Les premiers champs en haut nous renseignent sur les caractéristiques du groupe et le plafond général de la police. Les champs regroupés au milieu représentent les taux de remboursement propre à chaque prestation. Et enfin les champs en dessous présentent les sous-plafonds propres aux prestations plafonnées.

La prise d'écran qui va suivre nous montre la disposition de cette page :

The screenshot shows a software window titled "Taificateur" with two tabs: "Page1" and "Page2". The main content area features a background image of hands holding white paper cutouts of a family (a man, a woman, and a child). Overlaid on this image are several input fields and labels:

- Top left: "Effectif de l'entreprise" and "Plafond Général" with a text input field.
- Top right: "Nombre de conjoints" and "Nombre d'enfants" with text input fields.
- Center: "Taux de Remboursement" with a vertical stack of six text input fields.
- Left side (under a child cutout): "Soins Ambulatoires", "Radiologie", "Dentaire", "Hospitalisation", "Prothese dentaire", and "monture", each with a text input field.
- Right side (under a woman cutout): "Analyse", "Optique", "Pharmacie", "Maternité", "ODF", and "Naissance", each with a text input field.
- Bottom left: "Prestation Sous Plafonnée" with "Prothese dentaire" and "monture" below it, each with a text input field.
- Bottom center: "Sous Plafond" with a text input field.
- Bottom right: "Sous Plafond" with a text input field.
- Bottom right corner: A button labeled "Tarifier".

Figure 13: Fenêtre "Saisie des données"

Une fois les champs renseignés l'utilisateur n'a qu'à appuyer sur le bouton « Tarifier » afin de remplir les champs de la deuxième page.

✓ La deuxième page nous renseigne sur la prime pure calculée par type de prestation et type de bénéficiaire. Les champs sur la droite nous calculent la prime totale du groupe par type de bénéficiaire et le champ en dessous nous calcule la prime totale de l'entreprise tous types de bénéficiaire confondus.

	Assuré	Conjoint	Enfant
Soins Ambulatoires	<input type="text"/>	<input type="text"/>	<input type="text"/>
Radiologie	<input type="text"/>	<input type="text"/>	<input type="text"/>
Dentaire	<input type="text"/>	<input type="text"/>	<input type="text"/>
Hospitalisation	<input type="text"/>	<input type="text"/>	<input type="text"/>
Analyse	<input type="text"/>	<input type="text"/>	<input type="text"/>
Optique	<input type="text"/>	<input type="text"/>	<input type="text"/>
Pharmacie	<input type="text"/>	<input type="text"/>	<input type="text"/>
Monture	<input type="text"/>	<input type="text"/>	<input type="text"/>
Naissance (Assuré et Conjoint)	<input type="text"/>	<input type="text"/>	<input type="text"/>
Maternité (Assuré et Conjoint)	<input type="text"/>	<input type="text"/>	<input type="text"/>
Prothèse dentaire (Assuré et Conjoint)	<input type="text"/>	<input type="text"/>	<input type="text"/>
ODF (Enfant)			<input type="text"/>

Prime Totale

Assuré

Conjoint

Enfant

Prime Totale

réinitialiser

Figure 14: Fenêtre "Tarification"

Le Bouton « réinitialiser » permet de supprimer le contenu de tous les champs pour une nouvelle utilisation.

Conclusion :

Le tarif a priori -obtenu dans cette partie- donne des résultats généralement satisfaisants. Cependant, nous avons pu constater qu'il est difficile d'ajuster un modèle de manière fiable, c'est-à-dire par des validations statistiques, si les données ne sont pas suffisamment nombreuses.

Cependant, le MLG, dans le cas d'un nombre de données suffisant, permet de sélectionner les variables explicatives de la fréquence, d'une part, et des coûts moyens d'autre part. Ce ne sont pas forcément les mêmes. Ceci permet donc de déterminer un tarif plus proche des caractéristiques de l'assuré que la méthode directe.

L'algorithme de segmentation CHAID nous a aussi permis d'avoir des classes homogènes. Ces dernières ont donné lieu à une tarification minimisant au maximum la mutualisation, ce qui est contradictoire avec l'essence même du principe des Assurances.

Cette dernière question relève plusieurs questions.

Nous pousserons notre étude un peu plus loin et nous allons à travers la chapitre 5 étudier le comportement de la consommation médicale du portefeuille de la compagnie pour permettre à la compagnie d'identifier les profils types de consommateurs.

CHAPITRE 5 : ETUDE COMPORTEMENTALE DE LA CONSOMMATION MEDICALE

Introduction

L'étude comportementale de la consommation médicale de ce portefeuille d'assurés a pour principal objectif d'identifier les profils types de consommateurs parmi les adhérents du groupe. Ainsi, la méthode statistique utilisée est une méthode appelée HYBRID CLUSTERING ou classification mixte.

Cette méthode mixte, qui a été développée par Wong, est connue sous le nom de Hybrid Clustering. Ainsi, elle consiste à associer une méthode d'analyse en composantes principales ou d'analyse factorielle, selon le type de variables étudiées, à des méthodes de partitionnement « k-means », afin de réduire le volume de données à traiter, et de classification ascendante hiérarchique (CAH) pour obtenir la classification finale des assurés.

Dans cette partie nous présenterons à la fois la méthodologie de l'étude, précisant les variables retenues, les différentes étapes de l'analyse statistique et les résultats statistiques, c'est-à-dire une description de chaque profil de consommation obtenu.

I. Présentation des méthodes de classification :

I.1) La méthode de partitionnement « k-means » :

k-means est l'une des plus simples méthodes de classification dont la procédure suit un moyen simple et facile de partitionner un ensemble des objets sur un certain nombre de groupe (k groupes) fixé a priori. Ainsi, cette méthode vise à minimiser une fonction objectif, dans ce cas, c'est une fonction de l'erreur quadratique. La fonction objectif est :

$$\sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2$$

Où $\|x_i^{(j)} - c_j\|$ est une mesure de distance à choisir, entre un point de données de l'objet i $x_i^{(j)}$ et le centre du groupe j , c_j .

Le fil de l'exécution de l'algorithme "k-means" est le suivant :

- **Etape 0:** Spécifier le nombre de classes désiré,
- **Etape 1:** Spécifier k points qui serviront de centres pour les classes initiales,
- **Etape 2:** Affecter chaque point au centre le plus proche au sens d'une distance choisie au départ. Il en ressort un partitionnement en k classes,
- **Etape 3:** Recalculer le centre de chaque classe, les nouveaux centres n'étant pas nécessairement des points de l'ensemble à partitionner,
- **Etape 4:** Répéter les étapes 2 et 3 jusqu'à la stabilité des centres, c'est à dire des variations minimales et stables de la fonction objectif au fil des itérations.

I.2) La méthode de classification ascendante hiérarchique :

La classification ascendante hiérarchique est une méthode de classification utilisée pour regrouper des objets similaires « en groupes ». Dans un premier temps, on commence par affecter chaque objet à un groupe, de telle sorte qu'on obtient N groupes, l'étape suivante consiste à trouver les ensembles constitués de paires de groupes les plus proches et les fusionner en un seul groupe de taille N .

L'algorithme de cette méthode se présente comme suit :

- **Etape 1** : On considère initialement toutes les observations comme étant des groupes ne contenant qu'une seule observation (singleton).
- **Etape 2** : on calcule les distances euclidiennes entre les groupes

$$d(i, j) = \sqrt{|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2}$$

- **Etape 3** : On fusionne les groupes les plus proches, c'est-à-dire les groupes s'éloignant entre eux d'une distance minimale, ensuite on remplace ces derniers par une seule classe.
- **Etape 4** : on reprend l'algorithme à partir de l'étape 2 jusqu'à ce que nous disposions que d'un seul groupe, qui contient N observations de départ.

I.3) La classification mixte :

L'algorithme de cette méthode se déroule en deux étapes : l'ensemble des observations à grouper subit un partitionnement initial en utilisant la méthode « k-means » de telle façon à obtenir quelques dizaines de groupes homogènes, on procède ensuite à une classification hiérarchique sur les centres de ces groupes, dont le dendrogramme suggéra éventuellement le nombre de classes finales.

Ainsi, cette méthode permet de combiner les points forts de la méthode hiérarchique (CAH) et la méthode de partitionnement (k-means), la précision et l'optimalité de la première et la rapidité de la seconde.

I.4) Choix du nombre de groupes de la méthode CAH :

Plusieurs méthodes permettent de sélectionner le nombre de groupes de manière optimale.

✓ R- Squared

Le R-carré est un indice qui mesure l'homogénéité des groupes, il mesure la proportion de la variance expliquée par ses classes. Donc plus est proche de 1, les groupes sont plus homogènes.

$$R^2 = I_R / I_{total}$$

$$I_{total} = I_a + I_R$$

Où :

- I_R est l'inertie inter-classes
- I_a est l'inertie intra-classes

Ce dernier doit être le plus proche possible de 1 sans avoir trop de classes.

✓ Pseudo-F

Le pseudo-F est destiné à capturer la raideur du groupe, il est égal au rapport de la somme moyenne des carrés entre les groupes et la somme moyenne des carrés au sein du groupe un grand pseudo-F indique généralement un bon regroupement.

$$Pseudo - F = \frac{R^2}{(G - 1)} \bigg/ \frac{1 - R^2}{n - G}$$

Où

- G est le nombre de groupes,
- R^2 est la proportion de la variance expliquée par les classes.

✓ Le Clustering Critère cubique (CCC)

Le clustering Critère cubique a été développé par Sas comme une mesure comparative de la déviation des groupes de la distribution attendue si les points de données ont été obtenus à partir d'une distribution uniforme.

$$CCC = \log \left[\frac{1 - E(R^2)}{1 - R^2} \right] * K$$

Où K est le nombre de groupes.

Règle de décision :

- $CCC > 2$: Bonne classification
- $0 < CCC < 2$: Classification à vérifier
- $CCC < 0$: individus hors normes ou petites classes.

✓ R-carré semi-partiel (SPRSQ)

Le R^2 semi-partiel (SPRSQ) est une mesure de la perte d'inertie interclasse provoquée en regroupant 2 classes. Le but étant d'avoir une inertie interclasse maximale, on recherche un faible SPRSQ suivi d'un fort SPRSQ à l'agrégation suivante un pic pour k classes et un creux pour k+1 classes indique une bonne classification en k+1 classes.

II. Classification des assurés du portefeuille étudiée :

Préalablement à l'étape de classification, De nouvelles variables ont été créées à partir des variables initiales. Ainsi, pour chaque type de prestation a été créée une variable indiquant le pourcentage de la consommation de cette prestation par rapport à la charge totale dépensée par un individu.

Ainsi, nous retenons les variables suivantes pour la suite des traitements de la classification :

- Variables caractérisant le groupe : l'âge moyen, le taux de remboursement, le plafond général, l'effectif du groupe.
- Variables liées à la consommation : regroupent les variables indiquant la consommation médicale pour chaque type de prestation, ainsi que le montant total dépensé par chaque groupe.
- Variables liées au pourcentage de la consommation par prestation : regroupent les variables indiquant le pourcentage de la consommation médicale pour chaque type de prestation par rapport à la consommation totale.

II.1) Réduction et homogénéisation des données :

La première étape de l'HYBRID CLUSTERING consiste à réduire le nombre des variables du modèle en créant de nouvelles variables synthétiques, combinaisons des variables initiales afin de pouvoir faciliter l'interprétation des données et permettre de simplifier la création d'une classification et d'un partitionnement des données.

Ainsi, puisque les variables retenues sont quantitatives, on va procéder à une analyse en composantes principales visant à réduire les données.

Les variables utilisés à cette étape sont les variables représentant le pourcentage de la consommation pour une prestation donné par rapport à la consommation totale d'un individu.

Ainsi, la procédure PRINCOMP du logiciel SAS, fournit pour les composantes principales les valeurs propres suivantes.

Valeurs propres de la matrice de corrélation				
	Valeur propre	Différence	Proportion	Cumulé
1	1.96781313	0.24603533	0.1312	0.1312
2	1.72177780	0.29301041	0.1148	0.2460
3	1.42876739	0.13521838	0.0953	0.3412
4	1.29354901	0.08389828	0.0862	0.4275
5	1.20965072	0.14119486	0.0806	0.5081
6	1.06845586	0.04603038	0.0712	0.5793
7	1.02242549	0.03712048	0.0682	0.6475
8	0.98530501	0.07896321	0.0657	0.7132
9	0.90634180	0.07563607	0.0604	0.7736
10	0.83070573	0.03622626	0.0554	0.8290
11	0.79447947	0.10208537	0.0530	0.8820
12	0.69239410	0.15641574	0.0462	0.9281
13	0.53597836	0.11000554	0.0357	0.9638
14	0.42597282	0.30958950	0.0284	0.9922
15	0.11638332		0.0078	1.0000

Tableau 45-les valeurs propres des axes factoriels

Nous constatons que les 7 premiers axes factoriels permettent d'expliquer pratiquement 65% du modèle. Ainsi, on retient les 7 axes pour lesquelles la valeur propre est supérieure à 1.

II.2) Application de la méthode de partitionnement « k-means » :

Suite à l'analyse en composantes principales réalisée dans la précédente étape, on récupère les coordonnées des individus sur les 7 axes factoriels qui seront par la suite les variables de classification à cette étape.

La première chose à faire est de choisir le nombre de partitions que l'on veut. Pour cela on choisit de suivre le critère de Wong qui indique que pour un partitionnement de n individus, un nombre idéal de partitions est de $n^{0.3}$ partitions.

Dans notre cas, n=1870, nous allons donc partitionner nos individus en 10 partitions.

Pour réaliser une partition de type « k-means » des individus avec le logiciel SAS, on utilise la procédure FASTCLUS.

✓ La qualité de la classification « k-means »

Le tableau suivant résume les indicateurs de la classification « k-means » :

Pseudo F Statistic	247.06
Observed Over-All R-Squared	0.54451
Approximate Expected Over-All R-Squared	0.49984
Cubic Clustering Criterion	17.353

Tableau 46-les indicateurs de qualité de classification "k-means"

Les indicateurs sont généralement bons, ce qui ramène à dire que la première classification est suffisante pour continuer.

II.3) Application de la méthode de classification ascendante hiérarchique (CAH) :

Tout d’abord, on effectue une classification ascendante hiérarchique sur les centres de classes obtenus dans l’étape précédente suivie d’un découpage optimal du dendrogramme issu de la CAH qui fournit les classes finales.

✓ Le choix du nombre de classe

-La méthode de Ward : cette méthode consiste à agréger les deux classes de telle sorte que l’inertie intra-classe augmente le moins possible (et donc que l’inertie interclasse diminue le moins possible)

Ainsi, en choisissant la méthode de Ward, on agrège à chaque itération les classes dont l’agrégation fait perdre le moins d’inertie interclasse. Il s’agit donc d’une optimisation pas à pas, qui ne dépend pas d’un choix initial arbitraire. Ainsi, on élabore le dendrogramme des agrégations successives pour choisir le niveau de coupure selon Ward.

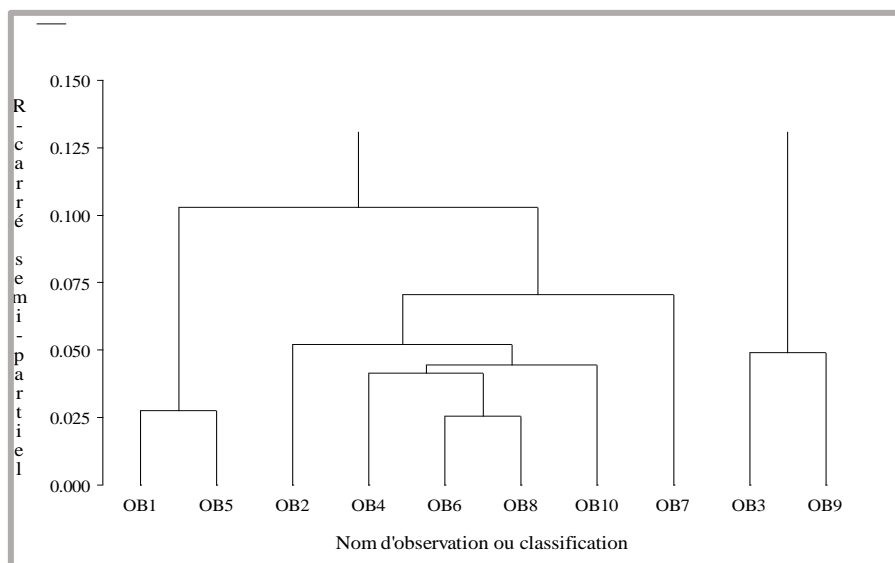


Figure 15- le dendrogramme des agrégations successives

Vu la forme de l'arbre, On décide de conserver 4 classes pour éviter la forte perte d'inertie interclasse en passant de 4 à 3 classes.

-Le Semi-Partiel : il mesure la perte d'inertie interclasse (ou de distance) provoquée en regroupant 2 classes. Le but étant d'avoir une inertie interclasse maximum, on recherche un faible SPRSQ suivi d'un fort SPRSQ à l'agrégation suivante : un pic pour k classes et un creux pour k+1 classes indique une bonne classification en k+1 classes.

‘ R carré semi partiel en fonction du nombre de classes’

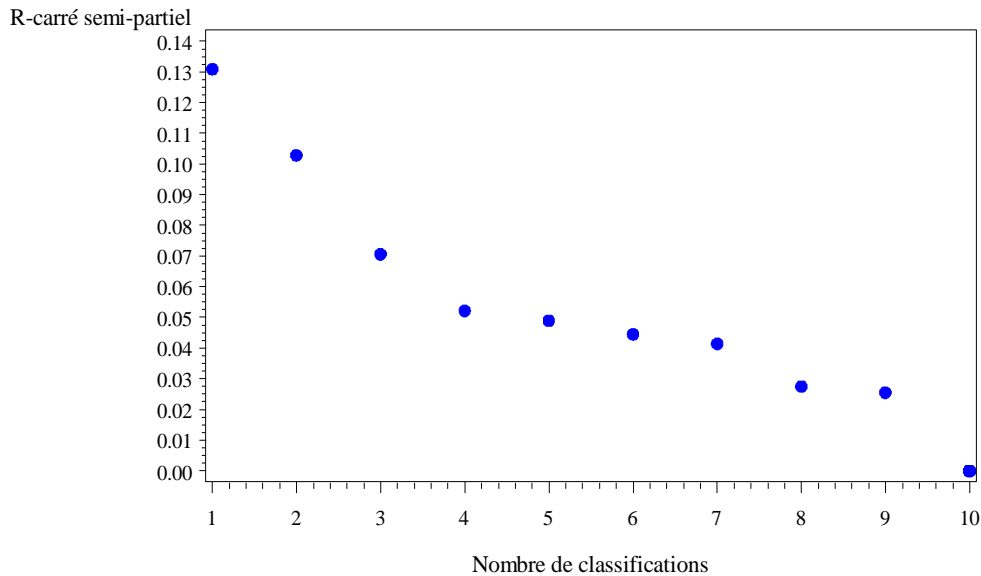


Figure 16-Graphe du R carré semi partiel en fonction du nombre de classes

D'après le graphe ci-dessus, nous pouvons constater que jusqu'à 4 classes (en lisant de droite à gauche), les regroupements de classe ne font que peu diminuer l'inertie interclasse (c'est-à-dire augmenter l'inertie intra-classe), mais que regrouper en 3 classes fait perdre énormément d'inertie interclasse. Donc nous retenons 4 classes.

✓ Description des classes obtenues

Pour rappel, l'arbre a été partitionné en 4 classes. Ainsi, la répartition du pourcentage des dépenses de chaque classe par prestation est la suivante :

	Classe 1	Classe 2	Classe 3	Classe 4
Proportion	46,47%	9,30%	32,46%	11,76%
Soins Ambulatoires	16.94 %	17.28 %	38.78 %	9.58 %
Analyse	10.52 %	5.95 %	5.85 %	4.46 %
Hospitalisation	9.8 %	1.8 %	5.29 %	51.02 %
Maternité	2.9 %	0.4 %	0.06 %	8.23 %
Optique	4.15%	25.00 %	3.04 %	1.82 %
Pharmacie	19.94 %	15.07%	34.2 %	10.12 %
Monture	2.23 %	15.72 %	2.26 %	1.18 %
Naissance	4.6 %	0.35 %	0.19 %	3.70 %

Prothèse	18.8 %	8.64 %	3.06 %	5.37 %
Radiologie	7.94 %	4.79 %	3.59 %	3.23 %
Dentaire	8.34 %	6.56 %	5.05 %	2.69 %
Consommation totale	206 612.51	31 847.28	46 483.43	137 500.34

Tableau 47-Répartition de la consommation par prestation (en % de la consommation totale de chaque classe)

La variable proportion représente le poids des effectifs de chaque classe dans la population totale. Les pourcentages indiqués pour chaque type de prestation représentent le pourcentage de ce poste dans la consommation totale pour chaque classe considérée

Par exemple, les dépenses en pharmacie représentent 19.94% de la consommation totale pour la classe 1 contre 34.2% pour la classe 3.

Les différentes prestations ne sont pas les seuls critères jouant un rôle dans ce partitionnement. Les caractéristiques démographiques et contractuelles des différentes classes d'individus sont les suivants :

		Classe 1	Classe 2	Classe 3	Classe 4
Proportion		46,47%	9,30%	32,46%	11,76%
Taux de remboursement	<80	12,66%	9,20%	12,19%	24,55%
	=80	44,76%	48,85%	52,72%	45,00%
] 80,85]	24,63%	20,69%	22,73%	19,55%
] 85,90]	10,82%	16,09%	6,92%	9,09%
	>90	7,13%	5,17%	5,44%	1,82%
Plafond Général	<25000	24,40%	24,14%	29,49%	25,91%
] 25000,35000]	17,03%	19,54%	11,20%	19,09%
] 35000,40000]	10,93%	8,62%	8,57%	10,45%
] 40000,50000]	28,65%	28,16%	36,90%	26,36%
] 50000,80000]	10,82%	14,94%	6,92%	7,27%
	>80000	8,17%	4,60%	6,92%	10,91%
Effectif du groupe	<=9	16,46%	30,46%	16,97%	10,45%
] 9,14]	11,51%	20,11%	10,54%	11,82%
] 14,31]	20,83%	21,26%	16,97%	25,00%
] 31,49]	10,36%	10,34%	9,72%	9,09%
] 49,70]	7,83%	5,75%	7,25%	6,36%
] 70,107]	8,17%	6,32%	14,17%	8,64%
] 107,190]	8,40%	1,72%	6,26%	9,55%
] 190,401]	8,52%	3,45%	5,60%	5,45%
	>401	7,94%	0,57%	12,52%	13,64%

		Classe 1	Classe 2	Classe 3	Classe 4
Proportion		46,47%	9,30%	32,46%	11,76%
lien	Assuré	54,32%	65,52%	16,47%	21,36%
	Conjoint	41,77%	21,84%	14,17%	65,91%
	Enfant	3,91%	12,64%	69,36%	12,73%
Classe âge	<=10	0,23%	0%	2,47%	0%
] 10,20]	0,23%	1,72%	6,43%	1,82%
] 20,30]	8,86%	16,09%	62,11%	16,36%
] 30,50]	87,46%	77,59%	27,18%	80,45%
	>50	3,22%	4,60%	1,81%	1,36%

Tableau 48- Caractéristiques sociodémographiques des classes d'individus

Le pourcentage indiqué pour chaque modalité dans une classe représente le pourcentage d'individus présentant cette caractéristique dans cette classe. Ainsi, 20.8% des groupes de la classe 1, ont un taux de remboursement égal à 80% tandis que cette même modalité représente 17.11% des groupes de la classe 3.

– **Classe n° 1**

La 1^{ère} classe regroupe 46.47% de la population étudiée. Le Montant de ses frais engagés est de 206 612 Dhs. Il s'agit de la classe qui a la consommation la plus grande. Elle regroupe ainsi presque le même pourcentage des assurés et des conjoints et enfants. Elle est constituée principalement des adhérents ayant un âge compris entre 30 et 50. De plus, elle regroupe les groupes d'assurés de faible effectif ainsi qu'un taux de remboursement égale à 80%.

La consommation médicale est concentrée essentiellement dans les prestations « Pharmacie » (19.94%) et « Prothèse » (18.8%).

– **Classe n°2**

Cette classe représente la plus faible proportion de la population étudiée (9.3%) dont 65% sont des assurés principaux moyennement jeune (âge entre 30 et 50). le montant de ses frais engagés est de 31 847 Dhs. Ainsi, nous retrouvons dans cette classe une grande part des groupes dont l'effectif est très faible (inférieur à 9) avec un plafond général en moyenne compris entre 40000 et 50000.

Par ailleurs, la consommation médicale est concentrée essentiellement dans la prestation « optique » (25%) suivi par celle des « soins ambulatoires » (17.28%).

– **Classe n°3**

La 3^{ème} classe regroupe 32% de la population étudiée. Le Montant de ses frais engagés est de 46 483 Dhs. Il s'agit d'une classe regroupant ainsi un mélange de groupes à effectifs différents, et est essentiellement constituée des enfants (69%) âgés de 20 à 30 ans avec un taux de remboursement majoritaire de 80%

La consommation médicale est concentrée essentiellement dans les prestations « Soins ambulatoires » (38.78%) et « Pharmacie » (34.2%).

– **Classe n°4**

Cette classe regroupe 51.02% de la population étudiée. Le Montant de ses frais engagés est de 137 500 Dhs. Il s'agit d'une classe regroupant des groupes faible effectifs (inférieur à 31), et est essentiellement constituée des Conjointes (66%) âgés de 30 à 50 ans avec un taux de remboursement majoritaire de 80%

La consommation médicale est concentrée essentiellement dans la prestation « Hospitalisation ».

Conclusion :

Une fois les primes calculées, l'analyse détaillé de la consommation des différentes prestations effectuée, l'entreprise sera amenée à calculer les provisions afin d'estimer les charges futures probables. Le calcul de ces provisions fera l'objet du chapitre suivant.

CHAPITRE 6 : PROVISIONNEMENT

Introduction

L'assurance a la spécialité d'avoir un cycle de production inversé : le coût d'un produit n'est connu qu'après sa vente. En effet, dans un premier temps, l'assureur reçoit une prime. Et dans un second temps, ce dernier devra payer à l'assuré un certain montant si un sinistre survient.

Les provisions Techniques représentent la grande partie du bilan d'une compagnie d'assurance 'Non Vie'. Une bonne approximation du niveau de provisions techniques, exigé en non vie, est vitale pour pouvoir surveiller la solvabilité par les régulateurs.

En outre, il faut souligner que la mauvaise estimation des fonds propres nécessaires, basée sur une évaluation erronée des Provisions Techniques, pourrait mener à des décisions peu appropriées avec l'état de santé financière d'un assureur.

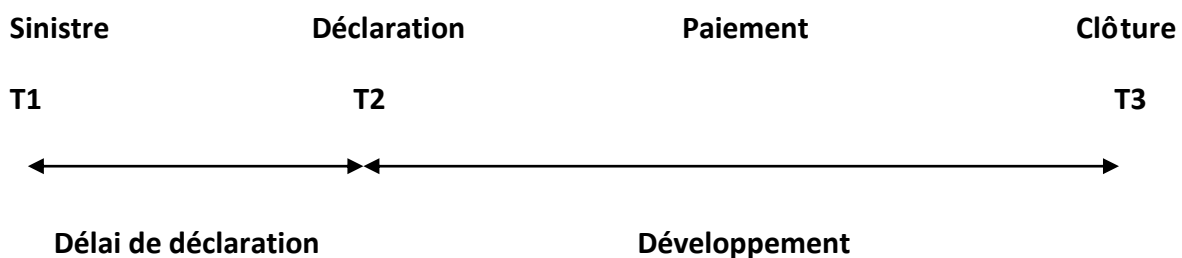
Dans ce qui suit nous allons nous intéresser à estimer les provisions de la branche maladie et plus particulièrement aux provisions pour sinistre à payer (PSAP) définies par l'arrêt réglementaire du 10 juin 1996 comme la valeur estimative des dépenses en frais et en principal, tant internes qu'externes nécessaire au règlement de tous les sinistres survenus à laquelle on ajoute les frais de gestion des sinistres ainsi que la charge des sinistres tardifs.

Nous allons traiter dans un premier lieu les méthodes déterministes et dans un second lieu nous nous intéresserons aux méthodes stochastiques.

I. La vie dynamique de vie d'un sinistre :

A un moment dans le temps, un sinistre survient. Avec un certain délai, le sinistre est déclaré à la compagnie.

Le processus qui va de la déclaration à la clôture du sinistre d'appelle le développement. Il est décrit par un ensemble de dates et la nouvelle information associée à chaque date.



Le sinistre survient à la date T1, est déclaré à la date T2 et clôturé à la date T3. L'exercice de provisionnement consiste à estimer à une date donnée le montant des paiements restant à faire pour l'ensemble des sinistres survenus.

Données utilisée :

Pour chacun des sinistres, Nous disposons des données :

- La date de survenance i
- La date de règlement j
- Les incréments de paiements $y_{i,j}$, pour l'année de règlement j et pour l'année de survenance i .

Année de survenance	Développement				
	1	...	j	...	n
1	$y_{1,1}$...	$y_{1,j}$...	$y_{1,n}$
.	.				.
.	.				.
.	.				.
i	$y_{i,1}$		$y_{i,j}$		
.	.		.		
.	.		.		
.	.		.		
n	$y_{n,1}$				

Tableau 49-Triangle des règlements des paiements

Le provisionnement est un problème de prédiction, conditionnelle à l'information dont on dispose à la date n.

II. La Méthode de Chain Ladder :

Cette méthode compte parmi les plus populaires, car elle est facile à mettre en œuvre, et facile à comprendre. L'idée est que le déroulement des paiements est gouverné par des facteurs de déroulement (notés λ_j) qui ne dépendent que de l'année de déroulement : le modèle sous-jacent est alors de la forme :

$$C_{i,j} = \lambda_j C_{i,j-1}$$

Les paramètres intervenant dans la méthode Chain Ladder présentent l'avantage d'avoir une interprétation claire, et sont facilement estimable. L'inconvénient étant que cette estimation est relativement peu robuste. Cette méthode ne fait aucune hypothèse quant à la loi que peuvent suivre les coûts des sinistres, ou leur fréquence.

La méthode Chain Ladder est fondée sur l'utilisation de Link-ratios encore appelés coefficients de passage, coefficients de déroulement ou coefficients de développement, entre les différents années de développement.

Les hypothèses sous-jacentes sont :

- **Hypothèse1** : les années de survenance sont indépendantes entre elles.
- **Hypothèse2** : les années de développement sont les variables explicatives du comportement des sinistres futurs.

Cette méthode consiste à supposer que les $C_{i,j}$ sont liés par un modèle de la forme :

$$C_{i,k+1} = \lambda_k C_{i,k} \text{ Pour tout } i, k = 1, \dots, n$$

Les coefficients λ_k sont appelés link-ratios. Ils peuvent être estimés, à l'aide des observations, par le rapport des totaux aux éléments communs de deux colonnes successives, c'est-à-dire par :

$$\widehat{\lambda}_k = \frac{\sum_{i=1}^{n-k} C_{i,k+1}}{\sum_{i=1}^{n-k} C_{i,k}}$$

Et à partir de ces coefficients de passage, il est alors possible d'obtenir une estimation des montants de provisions, en considérant :

$$\widehat{C}_{i,j} = (\widehat{\lambda}_{n+1-i} \dots \widehat{\lambda}_{j-1}) C_{i,n+1-i}$$

III. Modèle de Mack :

Le modèle proposé par T.Mack est une vision stochastique du modèle Chain Ladder. En effet, est ajouté aux hypothèses de ce dernier des indicateurs de risque de prédiction déduits d'une hypothèse de volatilité au sein du triangle. Le modèle de Mack est non paramétrique au sens où aucune hypothèse de distribution n'est faite sur les composantes du triangle, il est également conditionnel au sens où les espérances sont prises connaissant les réalisations du triangle supérieur et il s'applique à des montants cumulés : nombres, paiements, charges sinistres,...

✓ Hypothèses du modèle :

– **Hypothèse 1 :** Indépendance des exercices d'origine

Les $(C_{i,j})_{j=1,\dots,n}$ et $(C_{i',j})_{j=1,\dots,n}$ sont, pour $i \neq i'$, indépendants

– **Hypothèse 2 :** Pour $j = 1, \dots, n - 1$, il existe un paramètre f_j tel que :

$$\mathbb{E}[C_{i,j+1}/C_{i,1}, \dots, C_{i,j}] = f_j * C_{i,j}$$

– **Hypothèse 3 :** Pour $j = 1, \dots, n - 1$, il existe un paramètre σ_j^2 tel que :

$$Var[C_{i,j+1}/C_{i,1}, \dots, C_{i,j}] = \sigma_j^2 * C_{i,j}$$

Sous ces hypothèses Mack a montré que ce modèle stochastique permet d'obtenir exactement les mêmes montants de provision que la méthode standard de Chain-Ladder, qui est la référence en matière de provisionnement.

✓ Estimation des paramètres :

On introduit les termes $\lambda_0, \dots, \lambda_j$ qui vérifient : $\forall k \in \{1, \dots, J - 1\}$

$$\begin{cases} \lambda_k = \frac{\lambda_{k+1}}{f_{k+1}} \\ \lambda_j = 1 \end{cases}$$

Ainsi, $\forall i \in \{1, \dots, I\}, k \in \{1, \dots, J\}$:

$$E\left(\frac{C_{i,k}}{\lambda_k} \middle| C_{i,1}, \dots, C_{i,k-1}\right) = \frac{C_{i,k-1}}{\lambda_{k-1}}$$

On déduit l'expression suivante :

$$E(C_{i,k+1} | C_{i,1} \dots, C_{i,k}) = C_{i,I-i} \prod_{i=I-i+1}^k f_i$$

Les f_i sont estimés par la méthode Chain-Ladder, et on retrouve ainsi les estimateurs suivants :

$$\hat{f}_k = \frac{\sum_{i=0}^{I-k} C_{i,k+1}}{\sum_{i=0}^{I-k} C_{i,k}}, \forall k \geq I - i + 1$$

On peut donc définir l'estimateur de la charge $\hat{C}_{i,k}$ par :

$$\hat{C}_{i,k} = C_{i,I-i} \prod_{i=I-i+1}^k \hat{f}_k$$

IV. Modèle de Vyllder et de Christophides :

De Vyllder a été un des premiers modèles économétriques de provisionnement. On suppose que :

$$C_{i,j} \sim N(\alpha_i * \beta_j, \sigma^2), \forall i, j.$$

On peut estimer les paramètres de la loi par moindres carrés, on obtient ainsi :

$$\hat{\alpha}_i = \frac{\sum_j C_{i,j} \hat{\beta}_j}{\sum_j \hat{\beta}_j^2} \text{ et } \hat{\beta}_j = \frac{\sum_i C_{i,j} \hat{\alpha}_i}{\sum_i \hat{\alpha}_i^2},$$

Pour implémenter le code sur R, on a effectué la transformation suivante :

$$\log(C_{i,j}) \sim N(a_i + b_j, \sigma^2), \forall i, j.$$

Le coefficient a_i représente l'effet de la ligne i et b_j l'effet de la colonne j .

Après régression on estime les paramètres du modèle et ainsi on retrouve les $\hat{C}_{i,j} = \text{Exp}(\hat{a}_i + \hat{b}_j)$, cependant cet estimateur est biaisé et doit être corrigé du biais en posant

$$\hat{C}_{i,j} = \text{exp}(\hat{a}_i + \hat{b}_j + \hat{\sigma}^2 / 2).$$

V. Bootstrap:

La méthode Bootstrap est une méthode de simulation, alternative aux méthodes déterministes, qui peut donner des résultats satisfaisants si l'on dispose de peu de d'observation.

La technique de Bootstrap consiste à partir d'un échantillon originel unique, à créer par tirage aléatoire avec remise de nouveaux échantillons. Ces nouveaux échantillons sont constitués de « pseudo données » ayant la même loi sous-jacente. La répétition du ré-échantillonnage à partir de données « sources » permet d'estimer la variabilité d'un paramètre déterminé. Ce ré-échantillonnage est appliqué aux résidus du modèle.

Donc l'étape suivante consiste à définir les résidus suivants :

$$r_{ij} = \frac{C_{ij} - \hat{m}_{ij}}{\sqrt{\hat{m}_{ij}}},$$

- C_{ij} : le montant des paiements
- \hat{m}_{ij} : le montant des paiements estimés en utilisant la méthode de Chain Ladder rétrospective.

Toute fois ces résidus ont besoin d’être ajustés afin d’avoir une variance unitaire. On considère alors classiquement :

$$r_{ij} = \sqrt{\frac{n}{n - k}} \frac{C_{ij} - \hat{m}_{ij}}{\sqrt{\hat{m}_{ij}}}$$

Où k est le nombre de paramètres estimés dans le modèle.

Pour générer ces erreurs, la méthode la plus usuelle est d’utiliser une simulation non paramétrique, c'est-à-dire que l’on va bootstrapper les résidus parmi les pseudos-résidus obtenus.

Une fois simulé un pseudo-triangle de liquidation, on prédit un montant de provision (par exemple via la méthode Chain-Ladder). La variance des provisions correspond à l’erreur d’estimations.

VI. GLM :

Le modèle linéaire généralisé appliqué au provisionnement cherche à expliquer la consommation par l’année de survenance i et l’année de règlement j. Ainsi nos variables explicatives sont l’année de survenance et l’année de règlement.

La théorie reste la même et est expliquée ci-haut au chapitre 4.

VII. Application des méthodes de provisionnement :

✓ *Chain Ladder* :

La méthode Chain-Ladder a pour but de compléter le triangle de liquidation en estimant les IBNR. En appliquant cette méthode sur notre triangle nous retrouvons les résultats suivants :

Link ratio	1	2	3	4
Valeur	1.170003	1.002500	1.000184	1.000047

Tableau 50-Facteurs de développement de la méthode Chain ladder

Nous constatons que les facteurs de développement sont stables dans le temps. Nous remarquons aussi que seul le premier facteur de développement est plus ou moins supérieur à 1 car l’assurance maladie est une branche à développement rapide.

Nous calculons ensuite la charge ultime et les réserves pour chaque année :

Année	2010	2011	2012	2013	2014
Charge ultime	205 021 538	213 494 276	197 912 936	198 085 546	197 975 360
Réserve	0	10 091	45 770	539 578	29 226 947

Tableau 51 : Réserves et écarts types estimés par la méthode Chain Ladder (en Dhs)

La réserve totale pour les cinq années est 29 822 388 Dhs.

✓ *Mack :*

Année	2010	2011	2012	2013	2014
Tardifs	0	10 092	45 771	539 578	29 226 947
Ecart type de Mack	0	1 989	25 361	327 672	4 882 493

Tableau 52: Réserves et écarts types estimés par la méthode Mack (en Dhs)

Nous constatons que le modèle Mack donne les mêmes résultats que le modèle Chain Ladder et nous permet aussi d'avoir l'écart type et d'en déduire les intervalles de confiance. Ainsi la provision estimée par la méthode Mack est 29 822 388 Dhs.

✓ *De Vylder :*

Le tableau 53 nous indique sur la significativité de chaque modalité des deux variables explicatives.

Coefficient	Estimation	Erreur Standard	T-value	P-value
Intercept	***	***	2071.64	< 2 E-16
Col2	***	***	5.72	0.0012
Col3	***	***	-2.85	0.0288
Col4	***	***	-2.87	0.0264
Col5	***	***	4.10	0.0063
Ligne2	***	***	16.71	2.93 E-06
Ligne3	***	***	15.32	4.88 E-05
Ligne4	***	***	13.13	1.20 E-05
Ligne5	***	***	10.38	4.67 E-05

Tableau 53: Coefficients estimés par la méthode De Vylder (en Dhs)

Donc nous remarquons que tous les coefficients sont significatifs. Intéressons-nous au coefficient de détermination de notre modèle et à la statistique de Fisher du modèle afin de valider ce modèle.

Critère de validation	valeur
R ²	0.9909
R ² ajusté	0.9787
Résidus Standard	0.01332
Statistique de Fisher	81.53
P-value de F	1.484 E-05

Tableau 54-Critère de validation du modèle

Nous remarquons que tous les critères nous permettent de valider le modèle. En effet le R² ajusté proche de 1 nous renseigne sur la très bonne qualité d’ajustement et la p-value associée à la statistique de Fisher nous permet la prise de décision d’accepter le modèle.

Ainsi à travers les coefficients de régression estimés on retrouve le tableau de liquidation qui permet de soustraire les tardifs et les réserves suivantes:

Année	2010	2011	2012	2013	2014
Réserve	0	1 461 193	1 278 187	2 082 691	30 838 483

Tableau 55: Réserves et écarts types estimés par la méthode Devylder (en Dhs)

On constate que le modèle de Vylder et de Christophides ne donne pas les mêmes résultats que les deux modèles précédents à savoir Chain Ladder et Mack.

✓ *Bootstrap :*

Année	2010	2011	2012	2013	2014
Tardifs	0	6 935	46 229	529 886	29 296 281
Ecart type	0	131 571	205 571	489 882	3 972 781

Tableau 56: Réserves et écarts types estimés par la méthode Bootstrap (en Dhs)

Ainsi la provision estimée par la méthode Bootstrap est 29 879 331 Dhs.

✓ *GLM :*

Année	2010	2011	2012	2013	2014
Tardifs	0	9 452	58 943	511 758	29 139 535
Ecart type	0	5 779	28 356	263 352	17 659 219

Tableau 57: Réserves et écarts types estimés par la méthode GLM (en Dhs)

La réserve finale estimée par la méthode GLM est 29 719 689 Dhs.

✓ *Comparaison des méthodes Bootstrap, Mack et GLM*

– En termes d’estimation de la provision :

Le tableau 57 représente les réserves estimées par les trois méthodes Bootstrap, Mack et GLM.

Année De survenance	Bootstrap	Mack	GLM
2010	0	0	0
2011	6 935	10 092	9 452
2012	46 229	45 771	58 943
2013	529 886	539 578	511 758
2014	29 296 281	29 226 947	29 139 535
Provision totale	29 879 331	29 822 388	29 719 689

Tableau 58- Comparaison des méthodes pour les réserves estimées (en Dhs)

Le tableau suivant résume la différence d’estimation entre les trois méthodes :

Méthodes comparées	Différence
Bootstrap Vs Mack	56 943
Bosstrap Vs GLM	159 642
Mack Vs GLM	102 699

Tableau 59-différence d'estimation des réserves (en Dhs)

Nous remarquons que les méthodes qui donnent à peu près les mêmes résultats sont Bootstrap et Mack. Essayons à présent d'analyser la précision de chaque méthode en comparant leurs écarts types.

Année De survenance	Bootstrap	Mack	GLM
2010	0	0	0
2011	131 571	1 989	5 779
2012	205 571	25 361	28 356
2013	489 882	327 672	263 352
2014	3 972 781	4 882 493	17 659 219

Tableau 60: Comparaison des Réserves des différentes méthodes (en Dhs)

– En termes d'intervalle de confiance :

D'après ce tableau nous constatons que la méthode GLM est la moins précise puisqu'elle fournit des intervalles de confiance très larges donc elle n'est pas adéquate à notre tableau à l'instar de la méthode De Vylder.

Nous remarquons que la méthode Mack donne de meilleurs résultats en termes de précision car ces intervalles de confiance sont plus réduits et en l'occurrence plus précis. Donc cette méthode sera retenue pour estimer les provisions :

Année de Survenance	Borne inférieur	Charge ultime estimée	Borne Supérieur
2010	0	0	0
2011	8 103	10 092	12 081
2012	20 410	45 771	71 132
2013	211 906	539 578	867 250
2014	24 344 454	29 226 947	34 109 440
Total des provisions	24 923 419,85	29 822 388	34 721 356,2

Tableau 61 : Estimation et intervalle de confiance des réserves par la méthode Mack (en Dhs)

Conclusion :

A travers cette étude nous concluons que les méthodes de provisionnement ne donnent généralement pas les mêmes résultats en termes d'estimation des réserves et de précision. Ainsi le travail de l'actuaire est de trouver la méthode qui estime le mieux les provisions.

CONCLUSION GENERALE

Dans le contexte concurrentiel existant au sein du marché de l'assurance Santé, chaque compagnie cherche à obtenir le meilleur tarif qui permettra de séduire les assurés et de couvrir ses engagements. Cette recherche est principalement axée sur la sophistication des modèles de tarification.

Vu que l'étude des groupes dans le cadre de l'assurance santé est un exercice difficile à cause de l'hétérogénéité des groupes face au risque, il est tout à fait louable d'avoir engagé une démarche de classification par l'algorithme CHAID et utilisé la tarification par prestation en s'appuyant sur le modèle linéaire généralisé. Cette classification s'est avérée fructueuse vu le seuil de significativité des variables et des modalités qu'elle a permis d'avoir car une mauvaise segmentation peut influencer l'importance et la pertinence d'une variable.

L'étude portant sur le comportement de la variable charge totale annuelle de l'année 2013 et 2014 vis-à-vis de la variable prestation nous a permis de confirmer la tarification adoptée au finale. La segmentation retenue permet à la compagnie de se prémunir contre les mauvais risques en dégageant des sous-groupes homogènes en termes de consommation.

A travers ce projet nous avons arrêté un tarif extrait de modèles statistiques et méthodes actuarielles permettant à Saham Assurance de se prémunir contre le risque croissant inhérent à la branche maladie.

Nous avons pu constater au travers de cette étude que l'ajustement d'un modèle précis et robuste est tout à fait compliqué car des variables tarifaires pertinentes n'ont pas été fournies par la compagnie. En effet, on pourrait imaginer d'autres variables qui expliqueraient les dépenses de santé à savoir la masse salariale et l'âge moyen du groupe.

Ensuite dans un souci de perfection, une étude a été réalisée dans le chapitre 5 pour permettre à la compagnie de connaître la typologie du risque à laquelle elle fait face.

Finalement nous avons estimés les provisions pour sinistre à payer grâce à la comparaison des résultats obtenus en appliquant différentes méthodes de provisionnement afin d'offrir à la compagnie une vision claire des dépenses futures.

L'outil de tarification tel que présenté au chapitre 4 a été fourni à la direction Actuariat pour être implémenté sur le réseau Saham Assurance.

J'invite ainsi d'autres personnes à compléter notre travail en étudiant et modélisant les ratios de Sinistralité afin de vérifier l'amélioration des ratios de sinistralité de la branche maladie à Saham Assurance. Tous les ingrédients sont ainsi présents à savoir les primes appliqués, les dépenses effectuées et les engagements futurs.

BIBLIOGRAPHIE

Livres

- Arthur CHARPENTIER, Mathématiques de l'assurance non-vie : Tome1, Principes fondamentaux de théorie du risque. Edition Economica, 2005.
- Arthur CHARPENTIER, Mathématiques de l'assurance non-vie : Tome2, Tarification et provisionnement. Edition Economica, 2005.
- Jean-Marie BOUROCHE : L'analyse des données. Edition Puf, 2010.
- McCullagh P. & Nelder J. ,Generalized linear models. Edition Chapman et Hall, 1989.
- Pierre THEROND, Assurance non vie : Théorie de la crédibilité, Edition 2004.

Rapports

- Rapport Annuel 2013 de l'AMO.
- Rapport Annuel 2013 de l'ANAM.
- Rapport d'Activité Annuel 2014 de Saham Assurance.
- Rapport « Secteur marocain des assurances Situation liminaire-Exercice 2013 ».

Webo-graphies

- [www.anam.ma/ RAPPORT ANNUEL GLOBALE](http://www.anam.ma/RAPPORT_ANNUEL_GLOBALE)
- www.ressources-actuarielles.net

Mémoires

- Nicolas FANGET, Analyse comportementale de la consommation médicale d'un portefeuille d'assurés, Mémoire présenté pour l'obtention du diplôme d'actuaire de Strasbourg, année universitaire 2009/2010.
- Jean-Baptiste, La tarification en santé, Mémoire présenté pour l'obtention du diplôme d'actuaire de l'Université de Lyon, année universitaire 2011/2012
- Myriam EL JERDY, Tarification des groupes en assurance santé, Mémoire présenté pour l'obtention du diplôme d'Actuaire à l'Université de Lyon.
- Julien ROLLET, L'effet modérateur du rest à charge sur les dépenses de santé, Mémoire présenté pour l'obtention du diplôme d'Actuaire de l'Université de Lyon, année universitaire 2010/2011.

Annexe 1

Présentation de Saham Assurance :

➤ Saham Groupe :

Créé en 1995, SAHAM Group est un acteur de référence dans plusieurs à toute valeur ajoutée : l'assurance, les centres d'activités de relations clients, la santé et l'immobilier.

Variable incubateur, le Groupe a structuré son développement autour de sa capacité à créer de la valeur. Il a ainsi procédé à l'acquisition réfléchie et ciblée de structures à potentiel qu'il a développées en améliorant leur rentabilité et en accroissant leur valeur.

Aujourd'hui, leader panafricain aux synergies multi-métiers, SAHAM est également fermement implanté au Moyen-Orient.

Il concrétise ainsi sa mission en répondant aux besoins essentiels du plus grand nombre, en protégeant et en assistant ses clients et en leur offrant l'accès à des soins et à des services de qualité.

Les activités de SAHAM Group s'organisent en quatre pôles :

- Le pôle assurance, sous l'identité de SAHAM Finances, regroupe les métiers d'assurance, d'assistance et de crédit à la consommation.
- Le pôle off-shoring est principalement dédié aux centres de relations clients.
- Le pôle immobilier pilote des projets de développement dans l'immobilier résidentiel et social au Maroc et en Afrique subsaharienne.
- Par ailleurs, investi d'un profond sens de responsabilités, le Groupe contribue à développer d'ambitieux projets sociaux dans les pays où il est implanté à travers la fondation SAHAM.

➤ Saham Assurance :

– Brève description de l'assurance au Maroc :

Le secteur marocain de l'assurance connaît une tendance haussière qui lui a permis de se hisser à la première place dans le monde arabe et au 47^{ème} rang mondial en termes de taux de pénétration de l'assurance. Aujourd'hui, le marché marocain est considéré comme « le plus mature » au Maghreb selon l'agence de notation internationale Standard and Poor's.

En termes de chiffre d'affaires, le Maroc maintient la 2^{ème} place sur le continent africain après l'Afrique du Sud, et le 3^{ème} rang dans le monde arabe après l'Arabie Saoudite et les Emirats Arabes Unis.

– Saham Assurance :

Saham assurance (ex-CNIA SAADA Assurance) a été créée en 1949 sous le nom de Compagnie Nord-africaine et Intercontinentale d'Assurance. Seize ans plus tard, elle devient filiale de la Caisse de dépôt et de gestion.

En 1997, l'État cède les parts qu'il détient dans cette compagnie à Arab Insurance Group, qui devient alors l'actionnaire majoritaire en détenant 67 % du capital.

En 2001, la Compagnie Nord-africaine et Intercontinentale d'Assurance change de nom pour devenir CNIA Assurance et entreprend un vaste chantier de restructuration.

Rachetée en 2005 par le Groupe Saham, CNIA Assurance aspire à devenir un des acteurs majeurs de l'assurance au Maroc, dans ce sens, elle rachètera en 2006 l'assureur marocain ESSAADA.

En juin 2009, CNIA Assurance fusionne avec les assurances ES SAADA et devient CNIA SAADA Assurance.

En mars 2014, la compagnie et l'ensemble de ses filiales en Afrique et au Moyen-Orient (excepté au Liban) changent de nom pour devenir Saham assurance.

L'assureur est présenté dans une vignette de pays à travers ses 44 filiales et emploie près de 2 000 collaborateurs.

Le secteur marocain de l'assurance connaît une tendance haussière qui lui a permis de se hisser à la première place dans le monde arabe et au 47^{ème} rang mondial en termes de taux de pénétration de l'assurance. Aujourd'hui, le marché marocain est considéré comme « le plus mature » au Maghreb selon l'agence de notation internationale Standard and Poor's.

Saham Assurance a réalisé des performances excellentes en 2014 puisque le chiffre d'affaires a augmenté, la croissance est portée par la branche non vie qui a vu son chiffre d'affaire grimper, ç cet effet la compagnie est leader de l'assurance automobile au Maroc. Ces performances s'expliquent par les nombreux ajustements effectués ces dernières années dans le cadre du projet d'assainissement et d'amélioration des résultats du secteur. Soulignons à ce titre la refonte du cadre juridique régissant le secteur, la libéralisation progressive des tarifs d'assurances, le renforcement du réseau de distribution des produits d'assurances et l'encouragement des opérations de regroupement et de fusion entre certaines entreprises d'assurances, afin d'améliorer leur assise financière.

– Direction Actuariat :

La direction Actuariat est l'architecte du risque au sein de l'entreprise. Son métier consiste à analyser le risque pour garantir la solvabilité de l'entreprise. Cette direction est responsable de l'analyse du risque pour garantir la solvabilité de l'entreprise et améliorer sa rentabilité. Ainsi, votre rôle est d'anticiper les conséquences de l'évolution des marchés économiques, des réglementations nationales et internationales. Vous êtes un acteur du changement, contribuant à créer une société plus sûre et mieux préparée.

En relation avec les équipes financières de l'entreprise, cette direction:

- ✓ Evalue les risques.
- ✓ Assure la conception et le développement de nouveaux produits.
- ✓ Pilote la tarification des risques.
- ✓ Réalise le calcul d'indicateurs de rentabilité ou du capital économique.
- ✓ Evalue les garanties commercialisées, par la modélisation des marchés financiers et des comportements clients et distributeurs.

Elle peut également être en charge de :

- ✓ L'inventaire et l'analyse des marges.
- ✓ La réassurance.
- ✓ L'adossement actif/passif.
- ✓ La mise en place de stratégies commerciales visant à améliorer la rentabilité du portefeuille.

Annexe 2

Loi 65.00 :

Cette loi constitue le fondement de la protection sociale en matière de santé. D'autres dispositifs et mesures d'accompagnement, y compris les décrets d'application de la présente loi, seront soumis aux partenaires sociaux et aux parties concernés avant leur édicition.

Dans ce cadre, la présente loi portant code de la couverture médicale de base institue :

- ✓ Une assurance maladie obligatoire de base (AMO) fondée sur les principes et les techniques de l'assurance sociale au profit des personnes exerçant une activité lucrative, des titulaires de pension, des anciens résistants et membres de l'armée de libération et des étudiants ;
- ✓ Un régime d'assistance médicale (RAMED) fondée sur les principes de l'assistance sociale et de la solidarité nationale au profit de la population démunie.

Ce code constitue le fondement de la protection sociale.

Article 114 de loi 65.00 :

A titre transitoire, et pendant une période de cinq ans renouvelable, à compter de la date de publication des décrets réglementaires de la présente loi, les organismes publics ou privés qui, à cette date, assurent à leurs salariés une couverture médicale à titre facultatif, soit au moyen de contrats groupe auprès de compagnies d'assurances, soit auprès de mutuelles, soit dans le cadre de caisses internes, peuvent continuer à assurer cette couverture, sous réserve de fournir la preuve de l'existence de cette couverture à la CNSS ou à la CNOPS, selon le cas, et ce, selon les modalités fixées par voie réglementaire. Dans ce cas, la couverture doit s'appliquer à l'ensemble des salariés y compris les salariés nouvellement recrutés au cours de la période de transition, ainsi que, le cas échéant, les titulaires de pensions bénéficiant déjà de cette couverture. A l'expiration de ce délai, les employeurs précités sont tenus de procéder à leur affiliation et à l'immatriculation de leurs salariés et, le cas échéant, des titulaires de pensions ayant bénéficié de la couverture précitée, au régime de l'assurance maladie obligatoire de base géré par la CNSS ou par la CNOPS, selon le cas. Dans tous les cas, sont conservés les droits acquis par les bénéficiaires desdits régimes, tant pour la partie prenant en charge les cotisations que pour le taux de couverture dont ils bénéficient.

Annexe 3

ANOVA à un facteur et test d'égalité des moyennes:

Pour réaliser l'analyse de l'égalité des moyennes, nous avons codifié nos variables. Ainsi chaque code présent dans les sorties sera référencié à une type de prestation.

Code	Prestation
1	Analyse (Biologie)
2	Dentaire
3	Hospitalisation
4	Monture
5	Maternité
6	Naissance
7	Optique
8	Pharmacie
9	Radiologie
10	Soins Ambulatoires
11	Parodontie et Prothèse dentaire (Assuré, Conjoint)
12	Parodontie, Prothèse dentaire et ODF(Enfant)

➤ Lien A :

Comparaisons multiples						
Variable dépendante: charge Tamhane						
I	J	Différence de moyennes (I-J)	Erreur standard	Signification	Intervalle de confiance à 95%	
					Borne inférieure	Borne supérieure
1		-1633,96507	2237,89576	1,000	-9167,7877	5899,8576
	3	-39990,41402	5999,17085	,000	-60229,9357	-19750,8923
	4	4976,6564	1494,22878	,056	-52,2515	10005,5647
	5	1861,08559	2137,22633	1,000	-5346,0493	9068,2205
	6	4183,51080	1573,64131	,411	-1123,7338	9490,7554
	7	-942,21872	1766,87088	1,000	-6888,6327	5004,1953
	8	-18355,15841	3102,63910	,000	-28799,5776	-7910,7392

	9	3023,75510	1465,74968	,929	-1909,1344	7956,6446
	10	-9718,55274	2494,32876	,007	-18113,5358	-1323,5697
	11	-17526,18709	3474,92813	,000	-29236,8516	-5815,5225
2	1	1633,96507	2237,89576	1,000	-5899,8576	9167,7877
	3	-38356,44895	6183,81850	,000	-59210,3027	-17502,5952
	4	6610,62166	2117,13958	,114	-518,6960	13739,9393
	5	3495,05066	2610,99536	1,000	-5299,5353	12289,6366
	6	5817,47587	2173,91521	,393	-1506,3194	13141,2712
	7	691,74634	2317,62660	1,000	-7109,8432	8493,3359
	8	-16721,19334	3446,14709	,000	-28318,2185	-5124,1682
	9	4657,72017	2097,13673	,830	-2404,5247	11719,9651
	10	-8084,58767	2910,53884	,306	-17878,4205	1709,2452
	11	-15892,22202	3784,79668	,002	-28639,1235	-3145,3205
	3	1	39990,41402	5999,17085	,000	19750,8923
2		38356,44895	6183,81850	,000	17502,5952	59210,3027
4		44967,07061	5955,17871	,000	24873,7195	65060,4217
5		41851,49961	6148,10298	,000	21116,0270	62586,9722
6		44173,92482	5975,59878	,000	24012,5222	64335,3274
7		39048,19529	6029,36701	,000	18708,2819	59388,1087
8		21635,25561	6546,58714	,063	-427,9806	43698,4918
9		43014,16912	5948,09683	,000	22944,3481	63083,9901
10		30271,86128	6281,17105	,000	9093,8852	51449,8374
11		22464,22693	6731,01473	,056	-216,7135	45145,1673
4		1	-4976,65659	1494,22878	,056	-10005,5647
	2	-6610,62166	2117,13958	,114	-13739,9393	518,6960
	3	-44967,07061	5955,17871	,000	-65060,4217	-24873,7195
	5	-3115,57100	2010,43254	1,000	-9900,3699	3669,2279
	6	-793,14579	1396,58496	1,000	-5510,0741	3923,7826
	7	-5918,87531	1611,19065	,016	-11342,5796	-495,1710
	8	-23331,81500	3016,69885	,000	-33488,5550	-13175,0750
	9	-1952,90148	1273,78362	1,000	-6239,8899	2334,0870
	10	-14695,2093	2386,58303	,000	-22729,1008	-6661,3179
	11	-22502,8436	3398,41553	,000	-33958,3356	-11047,3518
	5	1	-1861,08559	2137,22633	1,000	-9068,2205
2		-3495,05066	2610,99536	1,000	-12289,6366	5299,5353
3		-41851,49961	6148,10298	,000	-62586,9722	-21116,0270
4		3115,57100	2010,43254	1,000	-3669,2279	9900,3699
6		2322,42521	2070,13679	1,000	-4666,1581	9311,0085
7		-2803,30432	2220,57471	1,000	-10289,1309	4682,5223
8		-20216,24400	3381,63995	,000	-31599,8034	-8832,6846
9		1162,66951	1989,35705	1,000	-5551,9101	7877,2491

	10	-11579,63833	2833,86579	,003	-21120,9609	-2038,3158
	11	-19387,27268	3726,15685	,000	-31940,2165	-6834,3289
6	1	-4183,51080	1573,64131	,411	-9490,7554	1123,7338
	2	-5817,47587	2173,91521	,393	-13141,2712	1506,3194
	3	-44173,92482	5975,59878	,000	-64335,3274	-24012,5222
	4	793,14579	1396,58496	1,000	-3923,7826	5510,0741
	5	-2322,42521	2070,13679	1,000	-9311,0085	4666,1581
	7	-5125,72953	1685,10018	,147	-10806,8761	555,4170
	8	-22538,66921	3056,81195	,000	-32831,2243	-12246,1141
	9	-1159,75570	1366,07175	1,000	-5774,8362	3455,3248
	10	-13902,06354	2437,08963	,000	-22108,2667	-5695,8604
	11	-21709,69789	3434,07272	,000	-33285,1850	-10134,2107
	7	1	942,21872	1766,87088	1,000	-5004,1953
2		-691,74634	2317,62660	1,000	-8493,3359	7109,8432
3		-39048,19529	6029,36701	,000	-59388,1087	-18708,2819
4		5918,87531	1611,19065	,016	495,1710	11342,5796
5		2803,30432	2220,57471	1,000	-4682,5223	10289,1309
6		5125,72953	1685,10018	,147	-555,4170	10806,8761
8		-17412,93968	3160,63048	,000	-28051,7679	-6774,1115
9		3965,97383	1584,81474	,562	-1368,9561	9300,9038
10		-8776,33401	2566,10437	,041	-17412,3546	-140,3134
11		-11734,54126	5859,98910	,972	-33276,3033	9807,2208
8		1	-16583,96837	3526,80325	,000	-28467,8921
	2	18355,15841	3102,63910	,000	7910,7392	28799,5776
	3	16721,19334	3446,14709	,000	5124,1682	28318,2185
	4	-21635,25561	6546,58714	,063	-43698,4918	427,9806
	5	23331,81500	3016,69885	,000	13175,0750	33488,5550
	6	20216,24400	3381,63995	,000	8832,6846	31599,8034
	7	22538,66921	3056,81195	,000	12246,1141	32831,2243
	9	17412,93968	3160,63048	,000	6774,1115	28051,7679
	10	21378,91351	3002,69453	,000	11269,0720	31488,7550
	11	8636,60567	3617,93149	,679	-3536,6633	20809,8746
	9	1	828,97132	4352,34165	1,000	-13820,6642
2		-3023,75510	1465,74968	,929	-7956,6446	1909,1344
3		-4657,72017	2097,13673	,830	-11719,9651	2404,5247
4		-43014,16912	5948,09683	,000	-63083,9901	-22944,3481
5		1952,90148	1273,78362	1,000	-2334,0870	6239,8899
6		-1162,66951	1989,35705	1,000	-7877,2491	5551,9101
7		1159,75570	1366,07175	1,000	-3455,3248	5774,8362
8		-3965,97383	1584,81474	,562	-9300,9038	1368,9561
10		-21378,91351	3002,69453	,000	-31488,7550	-11269,0720

	11	-12742,30784	2368,85648	,000	-20716,7425	-4767,8732
10	1	-20549,94220	3385,99034	,000	-31963,9900	-9135,8944
	2	9718,55274	2494,32876	,007	1323,5697	18113,5358
	3	8084,58767	2910,53884	,306	-1709,2452	17878,4205
	4	-30271,86128	6281,17105	,000	-51449,8374	-9093,8852
	5	14695,20933	2386,58303	,000	6661,3179	22729,1008
	6	11579,63833	2833,86579	,003	2038,3158	21120,9609
	7	13902,06354	2437,08963	,000	5695,8604	22108,2667
	8	8776,33401	2566,10437	,041	140,3134	17412,3546
	9	-8636,60567	3617,93149	,679	-20809,8746	3536,6633
	11	12742,30784	2368,85648	,000	4767,8732	20716,7425
	11	1	-7807,63435	3941,85038	,961	-21080,0283
2		17526,18709	3474,92813	,000	5815,5225	29236,8516
3		15892,22202	3784,79668	,002	3145,3205	28639,1235
4		-22464,22693	6731,01473	,056	-45145,1673	216,7135
5		22502,84368	3398,41553	,000	11047,3518	33958,3356
6		19387,27268	3726,15685	,000	6834,3289	31940,2165
7		21709,69789	3434,07272	,000	10134,2107	33285,1850
8		16583,96837	3526,80325	,000	4700,0446	28467,8921
9		-828,97132	4352,34165	1,000	-15478,6069	13820,6642
10		20549,94220	3385,99034	,000	9135,8944	31963,9900
*. La différence moyenne est significative au niveau 0.05.						

➤ Lien C :

Test d'homogénéité des variances

Statistique de Levene	ddl1	ddl2	Signification
58,119	11	9067	,000

Tests d'égalité des moyennes

	Statistique	ddl1	ddl2	Sig.
Welch	38,512	11	572,392	,000
Brown-Forsythe	56,844	11	1613,368	,000

Comparaisons multiples						
Variable dépendante: charge						
Tamhane						
I	J	Différence de moyennes (I-J)	Erreur standard	Signification	Intervalle de confiance à 95%	
					Borne inférieure	Borne supérieure
1	2	271,78188	709,96463	1,000	-2118,3777	2661,9415
	3	-32083,83456	3397,50414	,000	-43552,6881	-20614,9810
	4	3328,79855	606,64569	,000	1285,7452	5371,8519
	5	-3441,31670	963,84950	,024	-6688,9453	-193,6881
	6	-5997,46873	1025,35201	,000	-9454,6357	-2540,3018
	7	1645,20742	699,54906	,714	-710,1679	4000,5828
	8	-8140,71217	1353,87594	,000	-12699,3918	-3582,0325
	9	1527,51071	691,50320	,839	-800,3182	3855,3396
	10	-3907,18492	926,60350	,002	-7025,9207	-788,4491
	11	-8525,41414	1244,89710	,000	-12724,6410	-4326,1873
	2	1	-271,78188	709,96463	1,000	-2661,9415
3		-32355,61643	3388,79377	,000	-43795,5804	-20915,6525
4		3057,01668	555,79480	,000	1184,6765	4929,3568
5		-3713,09858	932,68147	,005	-6856,6035	-569,5936
6		-6269,25061	996,11024	,000	-9628,9599	-2909,5414
7		1373,42554	655,94036	,914	-835,6461	3582,4972
8		-8412,49405	1331,86671	,000	-12897,5511	-3927,4370
9		1255,72883	647,35272	,972	-923,8636	3435,3213
10		-4178,96680	894,13808	,000	-7188,8439	-1169,0897
11		-8797,19601	1220,92492	,000	-12916,6737	-4677,7183
3		1	32083,83456	3397,50414	,000	20614,9810
	2	32355,61643	3388,79377	,000	20915,6525	43795,5804
	4	35412,63311	3368,66330	,000	24039,4228	46785,8435
	5	28642,51786	3450,92442	,000	16996,2884	40288,7473
	6	26086,36582	3468,60494	,000	14381,3380	37791,3937
	7	33729,04197	3386,62699	,000	22296,2624	45161,8215
	8	23943,12238	3579,50465	,000	11869,4599	36016,7849
	9	33611,34526	3384,97417	,000	22184,0517	45038,6388
	10	28176,64963	3440,70741	,000	16564,4207	39788,8786
	11	23558,42042	3539,72347	,000	11616,7850	35500,0558
	4	1	-3328,79855	606,64569	,000	-5371,8519
2		-3057,01668	555,79480	,000	-4929,3568	-1184,6765
3		-35412,63311	3368,66330	,000	-46785,8435	-24039,4228

	5	-6770,11525	856,65859	,000	-9659,8780	-3880,3525
	6	-9326,26729	925,31337	,000	-12450,3397	-6202,1949
	7	-1683,59114	542,42693	,121	-3511,4169	144,2347
	8	-11469,51073	1279,78050	,000	-15780,3848	-7158,6367
	9	-1801,28785	532,01011	,047	-3592,9493	-9,6264
	10	-7235,98348	814,52570	,000	-9978,8819	-4493,0851
	11	-11854,21269	1163,88439	,000	-15784,1484	-7924,2770
5	1	3441,31670	963,84950	,024	193,6881	6688,9453
	2	3713,09858	932,68147	,005	569,5936	6856,6035
	3	-28642,51786	3450,92442	,000	-40288,7473	-16996,2884
	4	6770,11525	856,65859	,000	3880,3525	9659,8780
	6	-2556,15203	1190,45861	,883	-6567,6305	1455,3265
	7	5086,52412	924,77773	,000	1969,3367	8203,7115
	8	-4699,39547	1482,84356	,098	-9691,2656	292,4746
	9	4968,82741	918,70650	,000	1872,1429	8065,5119
	10	-465,86822	1106,54372	1,000	-4191,7352	3259,9988
	11	-5084,09744	1384,05699	,016	-9749,0316	-419,1633
	6	1	5997,46873	1025,35201	,000	2540,3018
2		6269,25061	996,11024	,000	2909,5414	9628,9599
3		-26086,36582	3468,60494	,000	-37791,3937	-14381,3380
4		9326,26729	925,31337	,000	6202,1949	12450,3397
5		2556,15203	1190,45861	,883	-1455,3265	6567,6305
7		7642,67615	988,71367	,000	4307,5428	10977,8095
8		-2143,24344	1523,53730	1,000	-7272,5299	2986,0430
9		7524,97944	983,03739	,000	4208,9398	10841,0191
10		2090,28381	1160,50846	,993	-1818,9503	5999,5179
11		-2527,94541	1427,56949	,995	-7339,5101	2283,6193
7		1	-1645,20742	699,54906	,714	-4000,5828
	2	-1373,42554	655,94036	,914	-3582,4972	835,6461
	3	-33729,04197	3386,62699	,000	-45161,8215	-22296,2624
	4	1683,59114	542,42693	,121	-144,2347	3511,4169
	5	-5086,52412	924,77773	,000	-8203,7115	-1969,3367
	6	-7642,67615	988,71367	,000	-10977,8095	-4307,5428
	8	-9785,91959	1326,34386	,000	-14252,5309	-5319,3083
	9	-117,69671	635,91246	1,000	-2259,1052	2023,7118
	10	-5552,39234	885,89052	,000	-8534,7034	-2570,0812
	11	-10170,62155	1214,89785	,000	-14270,0915	-6071,1516
	8	1	8140,71217	1353,87594	,000	3582,0325
2		8412,49405	1331,86671	,000	3927,4370	12897,5511
3		-23943,12238	3579,50465	,000	-36016,7849	-11869,4599
4		11469,51073	1279,78050	,000	7158,6367	15780,3848
5		4699,39547	1482,84356	,098	-292,4746	9691,2656

	6	2143,24344	1523,53730	1,000	-2986,0430	7272,5299
	7	9785,91959	1326,34386	,000	5319,3083	14252,5309
	9	9668,22288	1322,11793	,000	5215,8227	14120,6231
	10	4233,52725	1458,90822	,220	-677,0929	9144,1474
	11	-384,70196	1679,16287	1,000	-6038,4879	5269,0840
9	1	-1527,51071	691,50320	,839	-3855,3396	800,3182
	2	-1255,72883	647,35272	,972	-3435,3213	923,8636
	3	-33611,34526	3384,97417	,000	-45038,6388	-22184,0517
	4	1801,28785	532,01011	,047	9,6264	3592,9493
	5	-4968,82741	918,70650	,000	-8065,5119	-1872,1429
	6	-7524,97944	983,03739	,000	-10841,0191	-4208,9398
	7	117,69671	635,91246	1,000	-2023,7118	2259,1052
	8	-9668,22288	1322,11793	,000	-14120,6231	-5215,8227
	10	-5434,69563	879,55091	,000	-8395,4886	-2473,9027
	11	-10052,92484	1210,28285	,000	-14136,9496	-5968,9000
10	1	3907,18492	926,60350	,002	788,4491	7025,9207
	2	4178,96680	894,13808	,000	1169,0897	7188,8439
	3	-28176,64963	3440,70741	,000	-39788,8786	-16564,4207
	4	7235,98348	814,52570	,000	4493,0851	9978,8819
	5	465,86822	1106,54372	1,000	-3259,9988	4191,7352
	6	-2090,28381	1160,50846	,993	-5999,5179	1818,9503
	7	5552,39234	885,89052	,000	2570,0812	8534,7034
	8	-4233,52725	1458,90822	,220	-9144,1474	677,0929
	9	5434,69563	879,55091	,000	2473,9027	8395,4886
	11	-4618,22921	1358,38210	,045	-9196,1318	-40,3266
11	1	8525,41414	1244,89710	,000	4326,1873	12724,6410
	2	8797,19601	1220,92492	,000	4677,7183	12916,6737
	3	-23558,42042	3539,72347	,000	-35500,0558	-11616,7850
	4	11854,21269	1163,88439	,000	7924,2770	15784,1484
	5	5084,09744	1384,05699	,016	419,1633	9749,0316
	6	2527,94541	1427,56949	,995	-2283,6193	7339,5101
	7	10170,62155	1214,89785	,000	6071,1516	14270,0915
	8	384,70196	1679,16287	1,000	-5269,0840	6038,4879
	9	10052,92484	1210,28285	,000	5968,9000	14136,9496
	10	4618,22921	1358,38210	,045	40,3266	9196,1318

*. La différence moyenne est significative au niveau 0.05.

➤ Lien E :

Test d'homogénéité des variances

Statistique de Levene	ddl1	ddl2	Signification
43,448	11	5550	,000

Tests d'égalité des moyennes

	Statistique	ddl1	ddl2	Sig.
Welch	47,951	11	186,026	,000
Brown-Forsythe	46,002	11	1576,710	,000

Comparaisons multiples							
Variable dépendante: charge							
Tamhane							
) v	(I v	(J)	Différence de moyennes (I-J)	Erreur standard	Signification	Intervalle de confiance à 95%	
						Borne inférieure	Borne supérieure
1	2	-1693,18592	579,28552	,209	-3646,4139	260,0421	
	3	-16571,21822	2506,46699	,000	-25057,2717	8085,1647	
	4	1328,64209	391,73447	,046	8,7213	2648,5629	
	5	3406,66921	339,59416	,000	2257,3167	4556,0217	
	6	1781,62754	873,64931	,993	-2547,1620	6110,4171	
	7	1040,74779	379,11479	,334	-236,6715	2318,1671	
	8	-8234,17756	979,59178	,000	-11536,4019	-4931,9532	
	9	1186,40977	384,15735	,127	-107,9218	2480,7413	
	10	-14386,77145	1489,21684	,000	-19407,0157	-9366,5272	
	12	-4065,25880	1023,22683	,006	-7532,7792	-597,7384	
2	1	1693,18592	579,28552	,209	-260,0421	3646,4139	
	3	-14878,03230	2536,71012	,000	-23463,5784	-6292,4862	
	4	3021,82801	553,15220	,000	1155,8325	4887,8235	
	5	5099,85512	517,53807	,000	3351,2100	6848,5002	
	6	3474,81346	956,96635	,183	-695,7083	7645,3352	
	7	2733,93370	544,28806	,000	897,5323	4570,3351	
	8	-6540,99165	1054,57173	,000	-10094,0702	-2987,9131	
	9	2879,59568	547,81231	,000	1031,4547	4727,7366	
	10	-12693,58554	1539,57406	,000	-17881,9004	-7505,2707	
	12	-2372,07288	1095,22354	,873	-6077,0001	1332,8544	

3	1	16571,21822	2506,46699	,000	8085,1647	25057,2717
	2	14878,03230	2536,71012	,000	6292,4862	23463,5784
	4	17899,86031	2500,55642	,000	9433,2304	26366,4902
	5	19977,88742	2492,92012	,000	11536,3339	28419,4410
	6	18352,84576	2619,65449	,000	9456,5442	27249,1473
	7	17611,96600	2498,61053	,000	9151,7302	26072,2018
	8	8337,04065	2656,86405	,112	-644,9458	17319,0271
	9	17757,62798	2499,38060	,000	9294,8621	26220,3939
	10	2184,44676	2883,93713	1,000	-7550,9001	11919,7936
	12	12505,95942	2673,26012	,000	3468,7197	21543,1991
4	1	-1328,64209	391,73447	,046	-2648,5629	-8,7213
	2	-3021,82801	553,15220	,000	-4887,8235	-1155,8325
	3	-17899,8603	2500,55642	,000	-26366,4902	-9433,2304
	5	2078,02712	292,79657	,000	1083,7408	3072,3135
	6	452,98545	856,54467	1,000	-3947,3776	4853,3485
	7	-287,89430	337,83674	1,000	-1426,4333	850,6447
	8	-9562,81966	964,36803	,000	-12814,2925	-6311,3468
	9	-142,23232	343,48582	1,000	-1299,7342	1015,2695
	10	-15715,41355	1479,24726	,000	-20702,4421	-10728,3850
	12	-5393,90089	1008,66188	,000	-8813,6809	-1974,1208
5	1	-3406,66921	339,59416	,000	-4556,0217	-2257,3167
	2	-5099,85512	517,53807	,000	-6848,5002	-3351,2100
	3	-19977,88742	2492,92012	,000	-28419,4410	-11536,3339
	4	-2078,02712	292,79657	,000	-3072,3135	-1083,7408
	6	-1625,04167	833,98865	,998	-6151,2773	2901,1940
	7	-2365,92142	275,68454	,000	-3303,6037	-1428,2392
	8	-11640,84677	944,39081	,000	-14825,9830	-8455,7106
	9	-2220,25944	282,57885	,000	-3180,5403	-1259,9786
	10	-17793,44066	1466,30173	,000	-22737,4095	-12849,4718
	12	-7471,92801	989,57925	,000	-10829,4651	-4114,3910
6	1	-1781,62754	873,64931	,993	-6110,4171	2547,1620
	2	-3474,81346	956,96635	,183	-7645,3352	695,7083
	3	-18352,84576	2619,65449	,000	-27249,1473	-9456,5442
	4	-452,98545	856,54467	1,000	-4853,3485	3947,3776
	5	1625,04167	833,98865	,998	-2901,1940	6151,2773
	7	-740,87975	850,84716	1,000	-5169,2467	3687,4872
	8	-10015,80511	1240,89199	,000	-14568,4188	-5463,1914
	9	-595,21777	853,10594	1,000	-5012,1995	3821,7640
	10	-16168,39900	1672,71626	,000	-21946,9916	-10389,8064
	12	-5846,88634	1275,61986	,003	-10494,6242	-1199,1485
7	1	-1040,74779	379,11479	,334	-2318,1671	236,6715
	2	-2733,93370	544,28806	,000	-4570,3351	-897,5323

	3	-17611,96600	2498,61053	,000	-26072,2018	-9151,7302
	4	287,89430	337,83674	1,000	-850,6447	1426,4333
	5	2365,92142	275,68454	,000	1428,2392	3303,6037
	6	740,87975	850,84716	1,000	-3687,4872	5169,2467
	8	-9274,92535	959,31113	,000	-12509,5404	-6040,3103
	9	145,66198	329,02073	1,000	-962,9861	1254,3100
	10	-15427,51924	1475,95548	,000	-20403,5808	-10451,4577
	12	-5106,00659	1003,82813	,000	-8509,9539	-1702,0593
8	1	8234,17756	979,59178	,000	4931,9532	11536,4019
	2	6540,99165	1054,57173	,000	2987,9131	10094,0702
	3	-8337,04065	2656,86405	,112	-17319,0271	644,9458
	4	9562,81966	964,36803	,000	6311,3468	12814,2925
	5	11640,84677	944,39081	,000	8455,7106	14825,9830
	6	10015,80511	1240,89199	,000	5463,1914	14568,4188
	7	9274,92535	959,31113	,000	6040,3103	12509,5404
	9	9420,58733	961,31508	,000	6179,2969	12661,8778
	10	-6152,59389	1730,40938	,025	-11979,7166	-325,4712
	12	4168,91876	1350,38623	,129	-385,6505	8723,4880
9	1	-1186,40977	384,15735	,127	-2480,7413	107,9218
	2	-2879,59568	547,81231	,000	-4727,7366	-1031,4547
	3	-17757,62798	2499,38060	,000	-26220,3939	-9294,8621
	4	142,23232	343,48582	1,000	-1015,2695	1299,7342
	5	2220,25944	282,57885	,000	1259,9786	3180,5403
	6	595,21777	853,10594	1,000	-3821,7640	5012,1995
	7	-145,66198	329,02073	1,000	-1254,3100	962,9861
	8	-9420,58733	961,31508	,000	-12661,8778	-6179,2969
	10	-15573,18122	1477,25875	,000	-20553,5834	-10592,7790
	12	-5251,66857	1005,74339	,000	-8661,8841	-1841,4530
10	1	14386,77145	1489,21684	,000	9366,5272	19407,0157
	2	12693,58554	1539,57406	,000	7505,2707	17881,9004
	3	-2184,44676	2883,93713	1,000	-11919,7936	7550,9001
	4	15715,41355	1479,24726	,000	10728,3850	20702,4421
	5	17793,44066	1466,30173	,000	12849,4718	22737,4095
	6	16168,39900	1672,71626	,000	10389,8064	21946,9916
	7	15427,51924	1475,95548	,000	10451,4577	20403,5808
	8	6152,59389	1730,40938	,025	325,4712	11979,7166
	9	15573,18122	1477,25875	,000	10592,7790	20553,5834
	12	10321,51265	1755,47991	,000	4406,7017	16236,3236
11	1	2497,03579	1443,81044	1,000	-3939,2782	8933,3498
	2	803,84988	1495,69718	1,000	-5600,7404	7208,4402
	3	-14074,18242	2860,75429	,000	-23893,5704	-4254,7945
	4	3825,67788	1433,52511	,751	-2622,5695	10273,9253

	5	5903,70500	1420,16285	,095	-563,3458	12370,7558
	6	4278,66333	1632,42196	,700	-2410,5641	10967,8908
	7	3537,78358	1430,12809	,865	-2914,8627	9990,4298
	8	-5737,14177	1691,48999	,155	-12299,5234	825,2398
	9	3683,44556	1431,47309	,811	-2767,4287	10134,3198
	10	-11889,73566	2029,60714	,000	-19191,6626	-4587,8087
*. La différence moyenne est significative au niveau 0.05.						

Annexe 4

Tarification par GLM :

Cette annexe présente les Deviances et significativités des variables tarifaires pour la modélisation du coût moyen et de la fréquence pour chaque prestation par type de bénéficiaire:

➤ Lien A :

Analyse :

• **Coût moyen :**

Loi		Gamma		Log-normal	
Criterion	DF	Value	ValueDF	Value	ValueDF
Scaled Deviance	219	0.0000	0.0000	1237.0000	1.0148
Scaled Pearson X2	219	6.7796312E16	5.5616334E13	1237.0000	1.0148

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		9.99	0.0005
che_pg	tran		9.57	0.0015
che_na	tran		8.46	0.018

• **Fréquence :**

Loi		Poisson		Binomale Négative	
Criterion	F	Value	ValueDF	Value	ValueDF
Scaled Deviance	219	2454 4.4663	20.13 49	1 228.6178	1.0079
Scaled Pearson X2	219	3032 0.0340	24.87 29	1 228.6178	1.0079

Statistiques LR pour Analyse de Type			
3			
ce	Sour	DL	Pr
		hi-2	> Khi-2
che_tr	tran	1.76	<.0001
che_pg	tran	0.75	0.0009
che_na	tran	99.82	<.0001

Hospitalisation :

- Coût moyen :

Loi		Gamma		Log-normal		
Criterion	F	Value	Value	D	V	ValueDF
		e	DF	F	alue	
Scaled Deviance	219	0.0000	0.0000	7	7	1.0246
				32	50.0000	
Scaled Pearson X2	219	6.7796312E16	5.5616334E13	7	7	1.0246
				32	50.0000	

Statistiques LR pour Analyse de Type			
3			
ce	Sour	DL	Pr > Khi-2
tran			<.000
che_tr		7.6	01
tran			0.006
che_pg		6.05	7
tran			0.040
che_na		6.15	2

- Fréquence :

Loi		Poisson		Binomiale Négative		
Criterion	F	Value	Value DF	D	V	ValueDF
Scaled Deviance		3058	4.178	7	6	0.8842
	32	.3957	1	32	47.2169	
Scaled Pearson X2		4147	5.665	7	8	1.1318
	32	.3436	8	32	28.4543	

Statistiques LR pour Analyse de Type			
3			
ce	Sour	DL	Pr > Khi-2
tran			<.0
che_tr		0.79	001
tran			0.0
che_pg		3.72	175
tran			<.0
che_na		62.94	001

Dentaire :• **Coût moyen :**

Loi		Gamma		Log-normal		
Criterion	F	Value	Value DF	D F	V alue	ValueDF
Scaled Deviance	084	0.0000	0.0000	1084	102.0000	1.0166
Scaled Pearson X2	084	775.4059	7374.3316	1084	102.0000	1.0166

Statistiques LR pour Analyse de Type 3				
Source	DL	hi-2	Pr > Khi-2	
che_tr		.93	0.0416	
che_na		1.12	0.049	
che_pg		3.21	<.0001	

• **Fréquence :**

Loi		Poisson		Binomiale Négative		
Criterion	F	Value	Value DF	D F	V alue	ValueDF
Scaled Deviance	084	1268.05317	11.6979	1084	076.3071	0.9929
Scaled Pearson X2	084	1321.94713	12.1951	1084	067.0162	0.9843

Statistiques LR pour Analyse de Type 3				
Source	DL	hi-2	Pr > Khi-2	
che_tr		5.25	<.0001	
che_na		73.85	<.0001	
che_pg		7.30	<.0001	

Maternité :

- **Coût Moyen :**

Loi		Gamma		Log-Normal		
Criterion	F	Value	Value DF	D	V	ValueDF
Scaled Deviance		0.00	0.000	4	4	1.0389
	63	00	0	63	81.0000	
Scaled Pearson X2		1736	3749.	4	4	1.0389
	63	150.5318	7852	63	81.0000	

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		.39	0.0521
che_na	tran		6.59	0.035
che_pg	tran		5.04	0.0102

- **Fréquence :**

Loi		Poisson		Binomiale Négative		
Criterion	F	Value	Value DF	D	V	ValueDF
Scaled Deviance		1006	21.73	4	4	1.0580
	63	1.9180	20	63	89.8770	
Scaled Pearson X2		1267	27.37	4	4	1.0119
	63	6.4067	88	63	68.4872	

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		3.41	0.0001
che_na	tran		64.62	<.0001
che_pg	tran		5.75	0.0001

Monture :

- **Coût moyen :**

Loi		Gamma		Log Normale		
Criterion	F	Value	Value	D	V	ValueDF
Scaled Deviance		0.00	0.000	1	1	1.0196
	021	00	0	021	041.0000	
Scaled Pearson		1666	1631.	1	1	1.0196
	X2 021	133.4626	8643	021	041.0000	

Statistiques LR pour Analyse de Type 3				
Source	DL	hi-2	Pr	> Khi-2
he_tr		8.45	<.00	01
he_na		9.94	<.00	01
he_pg		8.51	<.00	01
he_mt		4.95	<.00	01

- **Fréquence :**

Loi		Poisson		Binomiale Négative		
Criterion	F	Value	Value	D	V	ValueDF
Scaled Deviance		6683	6.546	1	8	0.8699
	021	.5647	1	021	88.1725	
Scaled Pearson		7853	7.692	1	8	0.8011
	X2 021	.7092	2	021	17.9154	

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
he_tr	tranc		0.87	<.0001
he_na	tranc		48.80	<.0001
he_pg	tranc		7.66	0.0034
he_mt	tranc		7.03	0.0002

Naissance :

- **Coût moyen :**

Loi		Gamma		Log-Normale			
Criterion	F	Value	Value DF	F	D	V	ValueDF
Scaled Deviance	07	289.1729	1.3970	2	07	2	1.0870
Scaled Pearson X2	07	56.7951	0.2744	2	07	2	1.0870
						25.0000	

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		.30	0.3167
che_na	tran		6.65	0.034
che_pg	tran		9.43	0.0016
it_ns	forfa		6.09	<.0001

- **Fréquence :**

Loi		Poisson		Binomiale Négative			
Criterion	F	Value	Value DF	F	D	V	ValueDF
Scaled Deviance	07	396.8417	1.9171	2	07	1	0.7311
						51.3399	

Loi		Poisson		Binomiale Négative		
Criterion	F	Value	Value DF	D	V	ValueDF
Scaled Pearson		421.	2.037	2	1	0.8192
X2	07	6716	1	07	69.5662	

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		.03	0.9855
che_na	tran		13.65	<.0001
che_pg	tran		7.08	0.0205
it_ns	forfa		8.56	<.0001

Pharmacie :

- Coût moyen :

Loi		Gamma		Log Normale		
Criterion	F	Value	Value DF	D	V	ValueDF
Scaled Deviance		0.00	0.000	1	1	1.0125
	437	00	0	437	455.0000	
Scaled Pearson		4.36	3.037	1	1	1.0125
X2	437	4986E19	5685E16	437	455.0000	

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		2.58	0.0135
che_na	tran		5.60	0.0485
che_pg	tran		6.88	<.0001

- Fréquence :

Loi		Poisson		Binomiale Négative		
Criterion	DF	Value	ValueDF	DF	Value	ValueDF
Scaled Deviance		1052	73.25	1	1	1.0872
	437	63.7105	24	437	562.3687	
Scaled Pearson X2		1264	87.96	1	1	0.9759
	437	07.0144	59	437	402.3059	

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		0.59	<.00 01
che_na	tran		05.14	<.00 01
che_pg	tran		8.39	0.00 25

Radiologie :

- **Coût moyen :**

Loi		Gamma		LN		
Criterion	F	Value	ValueDF	D	V	ValueDF
Scaled Deviance		1301	1.076	1	1	1.0149
	209	.4879	5	209	227.0000	
Scaled Pearson X2		1580	1.307	1	1	1.0149
	209	.6842	4	209	227.0000	

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	K hi-2	Pr > Khi-2
che_tr	tran		11.45	0.0 431
che_na	tran		17 .56	0.0 015
che_pg	tran		15 .62	0.0 481

- Fréquence :

Loi		Poisson		BN			
Criterion	F	Value	Value DF	F	D	Value V	ValueDF
Scaled Deviance		1345	1.113		1		0.9643
	209	.7635	1	209		165.8267	
Scaled Pearson		1434	1.186		1		0.9643
	X2 209	.3842	4	209		165.8267	

Statistiques LR pour Analyse de Type 3				
Source	DL	hi-2	Pr > Khi-2	
che_tr		0.57	<.0001	
che_na		4.26	<.0001	
che_pg		88.45	<.0001	

Soins Ambulatoires :

- Coût moyen :

Loi		Gamma		LN			
Criterion	F	Value	Value DF	F	D	Value V	ValueDF
Scaled Deviance		1513	1.053		1		1.0125
	347	.8330	5	347		455.0000	
Scaled Pearson		1513	1.053		1		1.0125
	X2 347	.8330	5	347		455.0000	

Statistiques LR pour Analyse de Type 3				
Source	DL	hi-2	Pr > Khi-2	
che_tr		5.11	0.0099	
che_na		8.47	<.0001	
che_pg		5.75	0.046	

- **Fréquence :**

Loi		Poisson		BN		
Criterion	F	Value	Value DF	D F	V alue	ValueDF
Scaled Deviance		1687	1.174	1	1	1.0843
	437	.6872	4	437	558.1767	
Scaled Pearson		0.00	0.000	1	1	1.0843
X2	437	00	0	437	558.1767	

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		2.61	0.0004
che_na	tran		3.47	<.0001
che_pg	tran		67.62	<.0001

Optique :

- **Coût moyen :**

Loi		Gamma		LN		
Criterion	F	Value	Value DF	D F	V alue	ValueDF
Scaled Deviance		1140	1.065	1	1	1.0168
	070	.0965	5	070	088.0000	
Scaled Pearson		9813	9.171	1	1	1.0168
X2	070	.6452	6	070	088.0000	

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		6.01	0.0068
che_na	tran		0.95	0.0271
che_pg	tran		2.75	0.0037

- Fréquence :

Loi		Poisson		BN		
Criterion	F	Value	Value DF	D F	V alue	ValueDF
Scaled Deviance		1687	1.174	1	1	1.0843
	437	.6872	4	437	558.1767	
Scaled Pearson		0.00	0.000	1	1	1.0843
X2	437	00	0	437	558.1767	

Statistiques LR pour Analyse de Type 3				
Source	DL	hi-2	Pr > Khi-2	
tranche		5.21	<.0001	
tranche		7.97	<.0001	
tranche		11.50	<.0001	

Prothèse Dentaire et Parodontie :

- Coût moyen :

Loi		Gamma		LN		
Criterion	F	Value	Value DF	D F	Value	Value DF
Scaled Deviance		10	1.195	8	901.	1.0227
	81	53.3578	6	81	0000	
Scaled Pearson		11	1.316	8	901.	1.0227
X2	81	60.0000	6	81	0000	

Statistiques LR pour Analyse de Type 3				
Source	DL	hi-2	Pr > Khi-2	
tranche		1.04	0.0506	
tranche		3.51	0.0001	
tranche		5.50	0.05	
tranche		7.03	0.0002	

- **Fréquence :**

Loi		Poisson		BN		
Criterion	F	Value	Value DF	F	Value	ValueDF
Scaled Deviance		755.	0.857	8	89	1.0183
	81	0516	0	81	7.1486	
Scaled Pearson X2		1480	1.680	8	89	1.0183
	81	.2562	2	81	7.1486	

Statistiques LR pour Analyse de Type 3				
	Source	DL	hi-2	Pr > Khi-2
tr	tranche_		5.21	<.0001
na	tranche_		.85	0.043
pg	tranche_		11.50	<.0001
_ppd	Tranche		1.04	0.004

➤ **Lien C :**

Analyse :

- **Coût moyen :**

Loi	Gamma			Log Normale	
	F	Value	Value DF	Value	Value DF
Scaled Deviance		101	1.	10	1.0
	000	0.00	0100	10.0000	100
Scaled Pearson X2		0.00	0.	10	1.0
	000	03	0000	10.0000	100

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		0.00	0.0005
che_pg	tran		7.60	0.0035

- **Fréquence :**

Loi	Poisson			Binomiale Négative	
	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	000	7804.7988	7.8048	1103.5261	1.1035
Scaled Pearson X2	000	13022.973	13.0230	1840.8688	1.8409

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		5.78	<.0001
che_pg	tran		7.81	0.0032

Hospitalisation:

- **Coût moyen :**

Loi	Gamma			Log Normale	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	60	910. 5595	1. 3796	67 0.0000	1.0 152
Scaled Pearson X2	60	218. 9759	0. 3318	67 0.0000	1.0 152

Statistiques LR pour Analyse de Type 3			
Source	DL	hi-2	Pr > Khi-2
tran			0.00
che_tr		4.33	63
tran			0.02
che_pg		2.55	80

- **Fréquence :**

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	60	188 9.2056	2. 8624	70 7.6630	1.0 722
Scaled Pearson X2	60	373 9.2905	5. 6656	13 76.7851	2.0 860

Statistiques LR pour Analyse de Type 3			
ce	Sour	DL	Pr > Khi-2
che_tr	tran		<.00 01
che_pg	tran		0.03 1
		3.47	
		2.29	

Monture :

- **Coût moyen :**

Loi	Gamma			LN	
	F	Val ue	V alueDF	Va lue	Va lueDF
Scaled Deviance	42	894. 6954	1. 3936	67 0.0000	1.0 152
Scaled Pearson X2	42	227. 4807	0. 3543	67 0.0000	1.0 152

Statistiques LR pour Analyse de Type 3			
ce	Sour	DL	Pr > Khi-2
he_tr	tranc		<.00 01
he_pg	tranc		0.04 3
he_mt	tranc		<.00 01
		59.07	
		1.46	
		2.84	

- **Fréquence :**

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	42	201 1.1449	3. 1326	64 2.3803	1.0 006
Scaled Pearson X2	42	288 1.9307	4. 4890	82 0.0877	1.2 774

Statistiques LR pour Analyse de Type 3			
Source	DL	hi-2	Pr > Khi-2
he_tr		2.44	<.0001
he_pg		9.43	0.0016
he_mt		4.5	<.0001

Maternité :

- Coût moyen :

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	54	0.00 00	0. 0000	66 4.0000	1.0 153

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Pearson X2	54	3.28137E15	5.017E12	664.0000	1.0153

Statistiques LR pour Analyse de Type 3				
Source	DL	hi-2	Pr > Khi-2	
che_tr		6.92	0.0020	
che_pg		0.76	0.0009	

- **Fréquence :**

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	54	8284.5321	12.6675	749.4667	1.1460
Scaled Pearson X2	54	11862.981	18.1391	1137.5572	1.7394

Statistiques LR pour Analyse de Type 3				
Source	DL	hi-2	Pr > Khi-2	
che_tr		0.49	0.0004	
che_pg		0.30	0.0011	

Naissance :

- **Coût moyen :**

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	44	748.5833	1.3761	55	4.0000 184
Scaled Pearson X2	44	143.0295	0.2629	55	4.0000 184

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		.05	0.9756
che_pg	tran		1.91	0.036
it_ns	forfa		4.80	<.0001

- **Fréquence :**

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	44	132.64446	2.4383	55	1.1003 131
Scaled Pearson X2	44	234.80929	4.3163	11	54.4169 221

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		.59	0.4505
che_pg	tran		1.03	0.0008
it_ns	forfa		.66	0.008

Pharmacie :• **Coût moyen :**

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	230	0.0000	0.0000	12.40.0000	1.0081
Scaled Pearson X2	230	108.51667	8.822.49	12.40.0000	1.0081

Statistiques LR pour Analyse de Type 3				
Source	DL	hi-2	Pr > Khi-2	
che_tr		1.34	0.023	
che_pg		5.63	0.008	

• **Fréquence :**

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	230	354.93.526	2.8.8565	13.33.1202	1.0838
Scaled Pearson X2	230	431.96.013	3.5.1187	15.03.8923	1.2227

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		4.36	<.0001
che_pg	tran		8.39	0.0025

Dentaire :• **Coût moyen :**

Loi	Gamma			LN	
	F	Value	V alueDF	Value	V alueDF
Scaled Deviance	03	121 9.0985	1. 5182	81 3.0000	1.0 125
Scaled Pearson X2	03	307 7.1699	3. 832092	81 3.0000	1.0 125

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		.56	0.0458
che_pg	tran		4.28	<.0001

• **Fréquence :**

Loi	Poisson			BN	
	F	Value	V alueDF	Value	V alueDF
Scaled Deviance	03	416 8.0751	5. 1906	83 7.9257	1.0 435
Scaled Pearson X2	03	498 3.7362	6. 2064	10 42.3091	1.2 980

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		4.25	<.0001
che_pg	tran		1.48	0.0426

Radiologie :

- **Coût moyen :**

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	V alueDF
Scaled Deviance	27	995.5239	1.0739	937.0000	1.0108
Scaled Pearson X2	27	101.05237	1.1470	937.0000	1.0108

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		.94	0.0413
che_pg	tran		2.83	0.025

- **Fréquence :**

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	V alueDF
Scaled Deviance	27	198.13571	2.1373	1019.3577	1.0996

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	V alueDF
Scaled Pearson X2	27	371 6.0000	4. 0086	99 0.3685	1.0 683

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		01.74	<.00 01
che_pg	tran		1.12	0.04 9

Soins Ambulatoires :

- Coût moyen :

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	V alueDF
Scaled Deviance	233	130 5.7629	1. 0590	12 43.0000	1.0 081
Scaled Pearson X2	233	156 8.2387	1. 2718	12 43.0000	1.0 081

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		7.54	<.00 01
che_pg	tran		9.97	<.00 01

- Fréquence :

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	233	156 5.6634	1. 2698	13 53.7647	1.0 979
Scaled Pearson X2	233	183 2.4846	1. 4862	12 11.1759	0.9 823

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		09.63	<.00 01
che_pg	tran		4.13	0.01 48

Optique :

- Coût moyen :

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	85	753. 0368	1. 0993	69 5.0000	1.0 146
Scaled Pearson X2	85	753. 0368	1. 0993	69 5.0000	1.0 146

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		8.51	<.00 01
che_pg	tran		2.68	<.00 01

- Fréquence :

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	85	135 7.8755	1. 9823	76 7.6062	1.1 206
Scaled Pearson X2	85	180 9.7015	2. 6419	75 8.2781	1.1 069

Statistiques LR pour Analyse de Type 3				
Source	DL	Chi-2	Pr > Chi-2	
Model			0.0018	
Residual		1.64	0.04	

Prothèse dentaire et parodontie :

- Coût moyen :

Loi	Gamma			LN	
Criterion	F	Value	ValueDF	Value	ValueDF
Scaled Deviance	72	0.00 00	0.00 0	584. 0000	1.02 10
Scaled Pearson X2	72	3.16 82 E15	5.53 88 E12	584. 0000	1.02 10

Statistiques LR pour Analyse de Type 3				
	Source	DL	hi-2	Pr > Khi-2
	_tr	tranche	.74	0.04 5
	_pg	tranche	2.13	0.03 3
	_ppd	Tranche	2.07	<.00 01

- **Fréquence :**

Loi	Poisson			BN	
Criterion	F	Val ue	V alueDF	Va lue	Va lueDF
Scaled Deviance	72	726. 3584	1. 2698	58 7.1924	1.0 266
Scaled Pearson X2	72	867. 6812	1. 5169	58 7.1924	1.0 266

Statistiques LR pour Analyse de Type 3				
	Source	DL	hi-2	Pr > Khi-2
	_tr	tranche	7.16	0.00 18
	_pg	tranche	1.64	0.04
	_ppd	Tranche	8.69	<.00 01

- **Lien E :**

- **Pharmacie :**

- **Coût moyen :**

Loi	Gamma			Log Normale	
Criterion	F	Value	ValueD F	V alue	ValueDF
Scaled Deviance	41	0.0000	0.0000	8 51.0000	1.0119
Scaled Pearson X2	41	6.7348 461E16	8.00814 05E13	8 51.0000	1.0119

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		5.28	<.00 01
che_pg	tran		0.65	<.00 01

- **Fréquence :**

Loi	Gamma			Log Normale	
Criterion	F	Val ue	V alueDF	Va lue	Va lueDF
Scaled Deviance	41	477 91.587	5 6.8271	11 77.8951	1.4 006
Scaled Pearson X2	41	111 982.23	1 33.1537	20 083.112	23. 8800

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		0.47	<.0001
che_pg	tran		4.62	<.0001

Analyse :

- **Coût moyen :**

Loi	Gamma			Log Normale	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	63	100 8.9883	1. 5219	67 3.0000	1.0 151
Scaled Pearson X2	63	295 010984	4 44963.7	67 3.0000	1.0 151

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		.71	0.0457
che_pg	tran		4.94	0.0106

- **Fréquence :**

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	63	393 1.1309	5. 9293	75 7.2564	1.1 422
Scaled Pearson X2	63	689 6.2482	1 0.4016	23 46.8198	3.5 397

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		6.26	<.0001
che_pg	tran		5.74	0.0001

Hospitalisation :

- **Coût moyen :**

Loi	Gamma			LN	
	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	99	526.4882	1.3195	40.90000	1.0251
Scaled Pearson X2	99	602.3837	1.5097	40.90000	1.0251

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		0.35	<0.0001
che_pg	tran		1.98	0.035

- **Fréquence :**

Loi	Poison			BN	
	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	99	927.5843	2.3248	44.30609	1.1104

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Pearson X2	99	203 7.0752	5. 1055	11 09.4840	2.7 807

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		7.95	<.0001
che_pg	tran		1.85	0.045

Monture :

- **Coût moyen :**

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	44	802. 3341	1. 4749	556 .0000	1.0 221
Scaled Pearson X2	44	194. 1258	0. 3568	556 .0000	1.0 221

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
he_tr	tranc		1.90	0.023
he_pg	tranc		2.29	0.031
he_mt	tranc		5.75	<.0001

- **Fréquence :**

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	44	234 2.2912	4. 3057	56 6.3943	1.0 412
Scaled Pearson X2	44	326 3.6740	5. 9994	71 9.7721	1.3 231

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
he_tr	tranc		7.57	0.00 15
he_pg	tranc		1.98	<.00 01
he_mt	tranc		3.39	<.00 01

Maternité :

- **Coût moyen :**

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	4	35.6 093	1. 4837	33. 0000	1.3 750
Scaled Pearson X2	4	34.2 822	1. 4284	33. 0000	1.3 750

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		0.54	<.0001
che_pg	tran		6.12	<.0001

- **Fréquence :**

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	V alueDF
Scaled Deviance	4	20.0831	0.8368	27.1849	1.1327
Scaled Pearson X2	4	27.4505	1.1438	38.9101	1.6213

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		6.17	0.0028
che_pg	tran		1.10	0.0495

Naissance :

- **Coût moyen :**

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	V alueDF
Scaled Deviance		8.0078	8.0078		

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	V alueDF
Scaled Pearson X2		7.9147	7.9147		

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		.00	.
che_pg	tran		8.56	<.0001
it_ns	forfa		4.44	<.0001

- **Fréquence :**

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	V alueDF
Scaled Deviance		0.4391	0.4391	2.2404	2.2404
Scaled Pearson X2		0.5247	0.5247	1.8099	1.8099

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		.00	.
che_pg	tran		2.6767	0.0267
it_ns	forfa		.4788	0.0088

Dentaire :

- Coût moyen :

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	54	845.7332	1.5266	564.0000	1.0181
Scaled Pearson X2	54	58130632239	104928939.06	564.0000	1.0181

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		.74	0.045
che_pg	tran		2.13	0.033

- Fréquence:

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	54	5248.7213	9.4742	645.2466	1.1647
Scaled Pearson X2	54	10850.2957	19.5854	3186.9425	5.7526

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		9.74	<.0001
che_pg	tran		6.10	<.0001

Radiologie :

- **Coût moyen :**

Loi	Gamma			LN	
	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	03	651.4743	1.0804	613.0000	1.0166
Scaled Pearson X2	03	651.4743	1.0804	613.0000	1.0166

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		0.63	0.031
che_pg	tran		1.58	0.041

- **Fréquence :**

Loi	Poisson			BN	
	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	03	997.4223	1.6541	676.0320	1.1211
Scaled Pearson X2	03	997.4223	1.6541	676.0320	1.1211

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		18.64	<.0001
che_pg	tran		1.12	0.049

Soins Ambulatoires :

- **Coût moyen :**

Loi	Gamma			LN	
	F	Value	V alueDF	Value	V alueDF
Scaled Deviance	31	964.7786	1.0374	94.10000	1.0107
Scaled Pearson X2	31	964.7786	1.0374	94.10000	1.0107

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		1.11	0.0003
che_pg	tran		8.25	<.0001

- **Fréquence :**

Loi	Poisson			BN	
	F	Value	V alueDF	Value	V alueDF
Scaled Deviance	31	277.54972	2.9812	12.764932	1.3711

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Pearson X2	31	151 9.9506	1. 6326	95 4.5543	1.0 253

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		7.81	<.0001
che_pg	tran		9.64	<.0001

Optique :

- Coût moyen :

Loi	Gamma			LN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	79	602. 0422	1. 0398	58 9.0000	1.0 173
Scaled Pearson X2	79	682. 8234	1. 1793	58 9.0000	1.0 173

Statistiques LR pour Analyse de Type 3				
ce	Sour	DL	hi-2	Pr > Khi-2
che_tr	tran		4.16	0.0068
che_pg	tran		0.51	0.0010

- Fréquence :

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	Va lueDF
Scaled Deviance	79	967. 3482	1. 6707	64 9.6115	1.1 220
Scaled Pearson X2	79	120 6.3491	2. 0835	64 9.6115	1.1 220

Statistiques LR pour Analyse de Type 3				
Source	DL	hi-2	Pr > Khi-2	
tranche		5.36	0.00 4	
tranche		7.95	0.00 3	

Odf, prothèse dentaire et parodontie :

- **Coût moyen :**

Loi	Gamma			LN	
Criterion	D F	V alue	Val ueDF	Value	Valu eDF
Scaled Deviance	2 74	30 3.5783	1.10 79	283. 6142	1.03 51
Scaled Pearson X2	2 74	30 3.5783	1.10 79	283. 6142	1.03 51

Statistiques LR pour Analyse de Type 3				
Source	DL	hi-2	Pr > Khi-2	
tranche		0.48	0.00 04	
tranche		1.07	0.05	
Tranche		7.45	<.00 01	

- **Fréquence :**

Loi	Poisson			BN	
Criterion	F	Value	V alueDF	Value	ValueDF
Scaled Deviance	74	358.9543	1.3100	289.3257	1.0559
Scaled Pearson X2	74	375.6481	1.3709	300.0102	1.0949

Statistiques LR pour Analyse de Type 3				
	Source	DL	hi-2	Pr > Khi-2
	_tr tranche		1.12	0.0003
	_pg tranche		6.31	0.006
	_ppd Tranche		8.91	<.0001

La significativité des classes :

Pour un souci de présentation chaque prestation sera représentée par un code :

Prestation	Code
Analyse 'Bilogie'	1
Dentaire	2
Hospitalisation	3
Radiologie	4

Prestation	Code
Soins Ambulatoires	5
Optique	6
Maternité	7

Lien A :

- Significativité des modalités des différentes variables pour la modélisation du Coût moyen :

Code de la prestation

Par amètre	1	2	3	4	5	6	7
Inte rcept	0, 0000	0. 0000	0. 0000	0. 0000	0. 0000	0. 0000	0. 0000
tran che_tr1	0, 0469	0, 0123	0, 0233	0. 0091	0. 0023	0. 0000	0, 0227
tran che_tr2	0, 0075	0, 0471	0, 0497	0. 0492	0. 0371	0. 0000	0, 0096
tran che_tr3	0, 0002	0, 0016	0, 0069	0. 0341	0. 0266	0. 0000	0, 0482
tran che_tr4	0, 0006	0, 0151	0, 0158	0. 0468	0. 0189	0. 0000	0, 0105
tran che_tr5							
tran che_pg1	0, 0386	0, 0297	0, 0379	0. 0339	0. 0041	0. 0375	0, 0328
tran che_pg2	0, 0323	0, 0089	0, 0053	0. 0224	0. 0028	0. 0168	0, 0094
tran che_pg3	0, 0202	0, 0159	0, 0298	0. 0291	0. 0225	0. 0062	0, 0147
tran che_pg4	0, 0479	0, 0414	0, 0287	0. 0316	0. 0212	0. 0320	0, 0137
tran che_pg5	0, 0341	0, 0065	0, 0235	0. 0278	0. 0186	0. 0507	0, 0201
tran che_pg6							
tran che_na1	0, 0113	0, 0001	0, 0485	0. 0114	0. 0142	0. 0191	0, 0004
tran che_na2	0, 0290	0, 0002	0, 0421	0. 0190	0. 0206	0. 0148	0, 0183
tran che_na3	0, 0412	0, 0001	0, 0109	0. 0218	0. 0418	0. 0183	0, 0252
tran che_na4	0, 0466	0, 0033	0, 0371	0. 0016	0. 0215	0. 0279	0, 0306

tran che_na5	0, 0429	0, 0356	0, 0363	0. 0052	0. 0194	0. 0135	0, 0475
tran che_na6	0, 0124	0, 0001	0, 0258	0. 0270	0. 0139	0. 0021	0, 0435
tran che_na7	0, 0444	0, 0337	0, 0129	0. 0101	0. 0173	0. 0016	0, 0254
tran che_na8	0, 0063	0, 0188	0, 0026	0. 0172	0. 0236	0. 0283	0, 0011
tran che_na9							
Scal e							

- **Significativité des modalités des différentes variables pour la modélisation de la fréquence :**

Par amètre	Code de la prestation						
	1	2	3	4	5	6	7
Inter cept	0. 0000	0. 0000	0. 0000	0. 0000	0. 0000	0. 0000	0. 0000
tran che_tr1	0, 0188	0. 0000	0. 0000	0. 0000	0. 0070	0. 0330	0, 0101
tran che_tr2	0, 0007	0, 0045	0, 0039	0. 0018	0. 0461	0. 0271	0. 0000
tran che_tr3	0, 0329	0, 0394	0, 0024	0. 0192	0. 0276	0. 0364	0, 0030
tran che_tr4	0, 0294	0, 0058	0, 0468	0. 0217	0. 0000	0. 0238	0, 0158
tran che_tr5							
tran che_pg1	0. 0000	0. 0000	0. 0000	0. 0011	0. 0234	0. 0014	0. 0000
tran che_pg2	0. 0000	0. 0000	0. 0000	0. 0485	0. 0039	0. 0048	0. 0000

tran che_pg3	0. 0000	0. 0000	0. 0000	0. 0186	0. 0035	0. 0352	0. 0000
tran che_pg4	0. 0000	0, 0001	0. 0000	0. 0339	0. 0116	0. 0491	0, 0012
tran che_pg5	0. 0000	0, 0039	0. 0000	0. 0054	0. 0367	0. 0125	0, 0010
tran che_pg6							
tran che_na1	0. 0000	0, 0270	0. 0000	0. 0000	0. 0000	0. 0254	0, 0018
tran che_na2	0, 0049	0, 0019	0, 0006	0. 0000	0. 0000	0. 0687	0, 0049
tran che_na3	0. 0000	0, 0417	0. 0000	0. 0068	0. 0000	0. 0000	0, 0313
tran che_na4	0, 0481	0, 0001	0, 0180	0. 0152	0. 0000	0. 0297	0, 0122
tran che_na5	0, 0399	0, 0002	0, 0356	0. 0000	0. 0000	0. 0067	0, 0334
tran che_na6	0, 0492	0, 0035	0, 0331	0. 0092	0. 0000	0. 0342	0, 0137
tran che_na7	0, 0246	0, 0031	0, 0062	0. 0067	0. 0000	0. 0015	0, 0055
tran che_na8	0, 0072	0, 0003	0, 0145	0. 0452	0. 0000	0. 0000	0, 0091
tran che_na9							
Scal e							

Monture :

Paramètre	P-value	
	Coût moyen	Fréquence
Intercept	0.0000	0.0000
tranche_tr1	0.0000	0,0190

tranche_tr2	0.0000	0,0098
tranche_tr3	0.0000	0,0306
tranche_tr4	0.0000	0,0219
tranche_tr5		
tranche_pg1	0,0013	0.0000
tranche_pg2	0,0218	0.0000
tranche_pg3	0,0186	0.0000
tranche_pg4	0,0108	0.0000
tranche_pg5	0,0389	0.0000
tranche_pg6		0.0000
tranche_mt1	0.0000	0.0000
tranche_mt2	0.0000	0.0000
tranche_mt3		
Scale		

Naissance :

Paamater	P-value	
	Coût moyen	Fréquence
Intercept	0.0000	0.0000
tranche_pg1	0,0211	0.0000
tranche_pg2	0,0215	0.0000
tranche_pg3	0,0196	0.0000
tranche_pg4	0,0482	0,0007
tranche_pg5	0,0075	0,0075
tranche_pg6		0,0011
tranche_na1	0,0211	0,0391
tranche_na2	0,0485	0,0035
tranche_na3	0,0339	
tranche_na4	0,0267	0,0244
tranche_na5	0,0289	0,0393
tranche_na6	0,0071	0,0206
tranche_na7	0,0025	0,0391

tranche_na8	0,0147	0,0241
tranche_na9		
forfait_ns1	0.0000	0,0149
forfait_ns2	0.0000	0,0086
forfait_ns3		
Scale		

Prothèse dentaire et parodontie :

Paamater	P-value	
	Coût moyen	Fréquence
Intercept	0.0000	0.0000
tranche_pg1	0.0026	0.0007
tranche_pg2	0.0326	0.0217
tranche_pg3	0.0118	0.0418
tranche_pg4	0.0438	0.0367
tranche_pg5	0.0220	0.0362
tranche_pg6		
tranche_tr1	0.0073	0.0083
tranche_tr2	0.0003	0.0248
tranche_tr3	0.0001	0.0277
tranche_tr4	0.0000	0.0327
tranche_tr5		
tranche_na1	0.0123	0.0000
tranche_na2	0.0094	0.0000
tranche_na3	0.0462	0.0000
tranche_na4	0.0361	0.0000
tranche_na5	0.0323	0.0000
tranche_na6	0.0303	0.0000
tranche_na7	0.0243	0.0000
tranche_na8	0.0223	0.0000
tranche_na9		
Tranche_ppd1	0.0000	0.0257
Tranche_ppd2	0.0066	0.0140
Tranche_ppd3		

Scale		
-------	--	--

Lien C :

- **Significativité des modalités des différentes variables pour la modélisation du Coût moyen :**

Par amètre	Code de la prestation						
	1	2	3	4	5	6	7
Inter cept	0. 0000	0. 0000	0, 0000	0. 0000	0. 0000	0. 0000	0. 0000
tran che_tr1	0, 0015	0, 0023	0, 0207	0. 0367	0. 0000	0. 0093	0, 0008
tran che_tr2	0, 0134	0, 0108	0, 0306	0. 0351	0. 0186	0. 0127	0, 0349
tran che_tr3	0, 0017	0, 0019	0, 0207	0. 0456	0. 0324	0. 0125	0, 0430
tran che_tr4	0, 0105	0, 0060	0, 0299	0. 0225	0. 0093	0. 0072	0, 0424
tran che_tr5							
tran che_pg1	0, 0400	0. 0000	0, 0446	0. 0121	0. 0000	0. 0248	0, 0405
tran che_pg2	0, 0010	0, 0017	0, 0010	0. 0068	0. 0000	0. 0334	0, 0490
tran che_pg3	0, 0327	0, 0172	0, 0094	0. 0234	0. 0015	0. 0275	0, 0289
tran che_pg4	0, 0223	0, 0257	0, 0223	0. 0236	0. 0000	0. 0193	0, 0422
tran che_pg5	0, 0220	0. 0000	0, 001	0. 0283	0. 0012	0. 0071	0, 0230
tran che_pg6							
Scal e							

- **Significativité des modalités des différentes variables pour la modélisation de la fréquence :**

Par amètre	Code de la prestation						
	1	2	3	4	5	6	7
Inter cept	0. 0000	0. 0000	0. 0000	0. 0000	0. 0084	0. 0000	0. 0000
tran che_tr1	0. 0000	0. 0000	0. 0000	0. 0000	0. 0000	0. 0011	0, 0070
tran che_tr2	0. 0000	0, 0072	0, 0143	0. 0038	0. 0013	0. 0493	0, 0331
tran che_tr3	0. 0000	0, 0112	0, 0025	0. 0021	0. 0024	0. 0468	0, 0297
tran che_tr4	0, 0054	0, 0174	0, 0065	0. 0223	0. 0479	0. 0275	0, 0361
tran che_tr5							
tran che_pg1	0, 0014	0, 0166	0, 0012	0. 0435	0. 0126	0. 00354	0, 0013
tran che_pg2	0, 0088	0, 0040	0, 0012	0. 0228	0. 0485	0. 0148	0, 0076
tran che_pg3	0, 0247	0, 0097	0, 0046	0. 0463	0. 0318	0. 0436	0, 0036
tran che_pg4	0, 0046	0, 0010	0, 0010	0. 0114	0. 0201	0. 0132	0, 0089
tran che_pg5	0, 0052	0, 0063	0, 0021	0. 0426	0. 0314	0. 0279	0, 0086
tran che_pg6							
Scal e							

Monture :

Paramètre	P-value	
	Coût moyen	Fréquence
Intercept	0.0000	0.0000
tranche_tr1	0,0236	0.0000
tranche_tr2	0,0277	0,0281
tranche_tr3	0,0131	0,0158
tranche_tr4	0,0051	0,0259
tranche_tr5		
tranche_pg1	0,0345	0,0443
tranche_pg2	0,0126	0,0213
tranche_pg3	0,0216	0,0049
tranche_pg4	0,0398	0,0224
tranche_pg5	0,0305	0,0269
tranche_pg6		
tranche_mt1	0,0034	0,0341
tranche_mt2	0.0000	0.0000
tranche_mt3		
Scale		

Naissance

Paamater	P-value	
	Coût moyen	Fréquence
Intercept	0.0000	0.0000
tranche_pg1	0,0298	0.0000
tranche_pg2	0,0383	0.0000
tranche_pg3	0,0368	0,0107
tranche_pg4	0,0459	0,0001
tranche_pg5	0,0395	0,0135
tranche_pg6		
forfait_ns1	0.0000	0,0202
forfait_ns2	0.0000	0,0015
forfait_ns3		
Scale		

Prothèse dentaire et parodontie :

Paamater	P-value	
	Coût moyen	Fréquence
Intercept	0.0000	0.0000
tranche_pg1	0.0415	0.0208
tranche_pg2	0.0381	0.0143
tranche_pg3	0.0414	0.0358
tranche_pg4	0.0195	0.0428
tranche_pg5	0.0019	0.0027
tranche_pg6		
tranche_tr1	0.0266	0.0009
tranche_tr2	0.0135	0.0479
tranche_tr3	0.0031	0.0418
tranche_tr4	0.0387	0.0281
tranche_tr5		
Tranche_ppd1	0.0000	0.0466
Tranche_ppd2	0.0060	0.0296
Tranche_ppd3		
Scale		

Lien E:

- **Significativité des modalités des différentes variables pour la modélisation du Coût moyen :**

Par amètre	Code de la prestation					
	1	2	3	4	5	6
Intercept	0. 0000	0. 0000	0. 0000	0. 0000	0. 0000	0. 0000
tranche_tr1	0, 0010	0, 0195	0, 0078	0. 0492	0. 0070	0. 0000
tranche_tr2	0, 0067	0, 0294	0, 0014	0. 0113	0. 0154	0. 0000

tran che_tr3	0, 0086	0, 0477	0, 0111	0. 0097	0. 0093	0. 0000
tran che_tr4	0, 0091	0, 0401	0, 0062	0. 0126	0. 0346	0. 0021
tran che_tr5						
tran che_pg1	0, 0031	0, 0249	0, 0024	0. 0365	0. 0146	0. 0418
tran che_pg2	0, 0043	0, 0091	0, 0074	0. 0492	0. 0018	0. 0343
tran che_pg3	0, 0049	0, 0227	0, 0093	0. 0106	0. 0031	0. 0296
tran che_pg4	0, 0015	0, 0452	0, 001	0. 0351	0. 0093	0. 0167
tran che_pg5	0, 0076	0, 0277	0, 0012	0. 0194	0. 0348	0. 0100
tran che_pg6						
Scal e						

- **Significativité des modalités des différentes variables pour la modélisation de la fréquence :**

Par amètre	Code de la prestation					
	1	2	3	4	5	6
Inter cept	0. 0000	0. 0000	0. 0000	0. 0000	0. 0268	0. 0000
tran che_tr1	0. 0000	0, 0058	0. 0000	0. 0000	0. 0000	0. 0070
tran che_tr2	0, 0109	0, 0168	0, 0047	0. 0180	0. 0492	0. 0301
tran che_tr3	0, 0087	0, 0012	0, 0105	0. 0382	0. 0264	0. 0336

tran che_tr4	0, 0015	0, 0116	0, 0011	0. 0348	0. 0406	0. 0248
tran che_tr5						
tran che_pg1	0, 0201	0, 0052	0, 0072	0. 0132	0. 0348	0. 0187
tran che_pg2	0, 0078	0, 0062	0, 0023	0. 0336	0. 0036	0. 0315
tran che_pg3	0, 0013	0, 0387	0, 0130	0. 0218	0. 0031	0. 0099
tran che_pg4	0. 0000	0, 0261	0, 0212	0. 0264	0. 0255	0. 0334
tran che_pg5	0, 0349	0, 0090	0, 0011	0. 0371	0. 0000	0. 0249
tran che_pg6						
Scal e						

Monture :

Paramètre	P-value	
	Coût moyen	Fréquence
Intercept	0.0000	0.0000
tranche_tr1	0,0014	0,0018
tranche_tr2	0,0051	0,0061
tranche_tr3	0,0101	0,0057
tranche_tr4	0,0156	0,0050
tranche_tr5		
tranche_pg1	0,0018	0,0063
tranche_pg2	0,0022	0,0052
tranche_pg3	0,0063	0,0048
tranche_pg4	0,0053	0,0292
tranche_pg5	0,0166	0,0041
tranche_pg6		

tranche_mt1	0,0068	0,0059
tranche_mt2	0,0053	0,0042
tranche_mt3		
Scale		

Odf, Parodontie et Prothèse dentaire :

Paramètre	P-value	
	Coût moyen	Fréquence
Intercept	0.0000	0.0000
tranche_tr1	0.0429	0.0007
tranche_tr2	0.0062	0.0398
tranche_tr3	0.0306	0.0446
tranche_tr4	0.0353	0.0042
tranche_tr5		
tranche_pg1	0.0001	0.0382
tranche_pg2	0.0009	0.0264
tranche_pg3	0.0387	0.0143
tranche_pg4	0.0174	0.0066
tranche_pg5	0.0117	0.0109
tranche_pg6		
tranche_odf1	0.0000	0.0469
tranche_odf2		
Scale		