ROYAUME DU MAROC
--*-*-*
HAUT COMMISSARIAT AU PLAN
--*-*-*-*-*-*
INSTITUT NATIONAL
DE STATISTIQUE ET D'ECONOMIE APPLIQUEE

INSEA

Projet de Fin d'Etudes

*****

Modélisation de la Prime d'une Assurance Maladie Complémentaire & Application sur Access

Préparé par : ***Mr Abdelrhani FATHALLAH***

Mr Alpha Mamoudou BALDE

Sous la direction de ***: M. Mohamed EL HAJ TIRARI (INSEA)***

M. Khalid EL ANBRI (MAMDA-MCMA)

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

Diplôme d'Ingénieur d'Etat

Filière : ACTUARIAT-FINANCE

Devant le jury composé de :

- ***M. Mohamed EL HAJ TIRARI (INSEA)***
- ***M. Abdelilah YAAKOUBD (INSEA)***
- ***M. Khalid EL ANBRI (MAMDA-MCMA)***

Juin 2009

Résumé

Il est certain que chacun de nous est exposé au risque d'avoir une maladie curable ou chronique nécessitant des soins, ou tout simplement, aux cas échéants, des actes préventifs, dont les dépenses deviennent plus en plus lourdes et difficiles à cerner même en présence d'une assurance maladie de base.

Une assurance maladie complémentaire (ou une complémentaire maladie) s'impose; elle a pour objet de prendre en charge tout ou une partie des dépenses qui ne sont pas couvertes par une assurance maladie de base, notamment par l'Assurance Maladie Obligatoire (AMO). Les soins, les séjours hospitaliers et les visite chez le médecin ou la pharmacie coûtent cher. Souscrire une complémentaire maladie permet donc d'être mieux remboursé de ses dépenses de santé. La CNSS et la CNOPS, ainsi que d'autres compagnies d'assurance vie, ne remboursent pas totalement ces frais, cette assurance santé complémentaire permet de compenser le Ticket Modérateur, qui n'est que la différence entre le coût réel d'une prestation de santé et de son remboursement par ses organismes.

Le présent travail a pour objectif d'établir une tarification à priori d'une assurance maladie complémentaire de type collectif, en procédant par deux approches différente: la première approche calcule la prime pure comme étant le produit de la fréquence de sinistralité, en l'ajustant par une loi de probabilité adéquate et une régression GLM sur les caractéristiques propres à chaque groupe d'assurés, avec le coût moyen des sinistres qu'on ajuste graphiquement par un Q-Q plot. Quant à la deuxième approche, elle élabore un modèle dans lequel la prime pure propre à chaque groupe est le produit de la fréquence de sinistres, ajustée par une régression logistique, avec la charge annuelle moyenne d'une police estimée par le Q-Q plot aussi.

La première partie de ce travail, édicte les principales caractéristiques du secteur marocain des assurances, plus précisément, elle donne des aperçus sur l'assurance maladie obligatoire (AMO) et l'assurance maladie complémentaire avant de donner une représentation de l'organisme d'accueil, à savoir la compagnie MAMDA-MCMA.

Dans la deuxième partie, on présentera la base de données qui a servi d'assise dans tout notre travail, puis on procédera à l'analyse descriptive de ses différentes composantes et à la segmentation du portefeuille santé de la compagnie, et aux descriptions les plus détaillées de variables tarifaires les plus pertinentes.

La troisième partie a pour objectif d'effectuer une tarification à priori de chaque contrat groupe. Il s'agit de déterminer, selon chacune de deux approches précitées, la prime pure de chaque police, c'est-à-dire le montant à hauteur duquel elle doit cotiser pour être indemnisé des coûts des sinistres causés au niveau de chaque poste de maladie .La cotisation totale du groupe serait la somme des primes dégagées dans tous les postes auxquels il a inscrit une assurance complémentaire.

Enfin, la quatrième partie, présente sommairement l'application informatique que nous avons conçu afin d'automatiser le fonctionnement des modèles élaborés et retenus dans notre travail. Cette application consiste, via un logiciel SGBD , de sélectionner les données nécessaires pour le calcul de la prime pure à partir de la base de données initiale et des résultats des modèles générés par les logiciels SPSS et EXCEL-STAT en vue d'élaborer un modèle conceptuel de données, avant d'implémenter celui-ci dans ACCESS.

Mots clés *: Assurance maladie obligatoire ; Assurance maladie complémentaire; Assurance de groupes; Charge moyenne d'une police; Coût annuel moyen d'un sinistre; Fréquence des sinistres; Merise; poste de maladie; Tarification à priori; Ticket Modérateur.*

Dédicace

A la mémoire de ma défunte mère,

Que Dieu l'accueille dans son paradis

A la mémoire de mon très cher père,

Que Dieu le Protège

A ma petite famille à Meknès,

A ma grande famille à Meknès, Rabat, Ben guérir et Casa,

A touts mes amis là où ils sont, à l'INSEA, partout au Maroc et à l'étranger,

A eux tous je dédis ce modeste travail…

Abdelrhani FATHALLAH

Dédicace

A la mémoire de mon petit frère Thierno.

Je dédie ce travail

En premier lieu à ma famille, mes chers parents pour leur soutien moral et matériel durant ces dix huit longues années d'études conduisant au diplôme d'ingénieur, un statut de promotion sociale de nos jours.

A mon ami nigérien doulaye Isaac pour notre amitié sincère durant nos deux années de deug à la faculté des sciences de Kenitra. a travers lui, j'ai compris qu'il vaut mieux avoir un vrai ami sur une base noble et immatérielle pendant un laps de temps que des centaines de relations humaines orientées objet pendant toute une vie.

A tous mes amis, tous azimuts qui m'ont soutenu pour que travail soit possible.

Merci encore.

Je vous serai reconnaissant à jamais.

Merci.

BALDE ALPHA MAMOUDOU.

Remerciements

Qu'il nous soit permis, au terme de ce travail, d'exprimer notre gratitude et nos vifs remerciements à notre directeur de mémoire Monsieur Mohamed El Haj Tirari. Qu'il trouve ici le témoignage de notre estime et de notre profonde reconnaissance pour sa disponibilité, ses conseils et sa compétence qu'elle a su nous prodiguer tout au long de notre projet.

Nos remerciements, les plus sincères, vont à notre encadrant monsieur Khalid El Anbri, directeur du département assurance vie de la MAMDA-MCMA pour son accueil, sa générosité, ses directives encourageantes et ses qualités humaines. Ainsi à monsieur Aadil El khaloui , responsable du département commercial et marketing à la direction vie ; et Qu'ils trouvent tous ici l'expression de notre profonde reconnaissance pour l'intérêt et l'effort qu'ils ont déployé afin de nous accompagner et nous soutenir durant notre période de stage.

Nous remercions également monsieur……. et monsieur…….. pour l'intérêt qu'ils ont porté à ce travail en acceptant de siéger parmi les membres du jury.

On saisit cette occasion pour exprimer notre gratitude et estime à notre corps professoral compétant de l'INSEA.

Notre sincère reconnaissance va à tous ceux qui, de près ou de loin, ont contribué à l'aboutissement et au bon déroulement de ce modeste travail.

Sommaire :

Liste des abréviations 9
Liste des tableaux 10
Liste des figures 13
Introduction 15

Partie I : Généralités & Présentation 17

Chapitre I : Les caractéristiques du secteur des assurances au Maroc 18

I .1 : les composantes du secteur marocain des assurances 18

I.2 Les apports et les nouveautés des années 2007 et 2008 18

I .3 Chiffres d'affaire des différentes branches du secteur marocain des assurances au titre de l'exercice 2007 : 19

Chapitre II: Présentation de l'organisme d'accueil MAMDA-MCMA 21

II.1 Fiche technique de la MAMDA-MCMA 21

II.2 Activités et performance de la MAMDA-MCMA 21

II.2.1 Activités de la MAMDA-MCMA 21

II.2.2 Expériences de la MAMDA-MCMA 23

II.2.3 Les Réassureurs de la MAMDA-MCMA 24

Chapitre III: Assurance maladie Obligatoire & Assurance maladie complémentaire. 24

III-1 Assurance maladie Obligatoire 24

III-1-1 Présentation : 24

III.1.2 Les prestations fournis par l'AMO : 25

III.2.1.1 Prestations couvertes et taux de Cotisation du secteur privé, CNSS : 27

2-1-1-1 Taux de cotisation de secteur privé : 29

2.1.1.2 Taux de couverture de secteur privé : 29

III.2.1.2 Prestations couvertes et taux de couverture du secteur public, CNOPS 29

III.2.1.2 .1 Taux de couverture de secteur public : 29

III.2.1.2 .2 Taux de cotisation de secteur public : 29

III.1.3 Les bénéficiaires : 30

III.1.3.1 Secteur public : 30

III-1-3-2 Secteur privé: 31

III.2 Assurance maladie complémentaire 32
III.2.1 Qu'est ce que l'assurance maladie complémentaire ? 32
III.2.2 Le remboursement du Ticket Modérateur : 32
III.2.3 La Tarification Nationale de Référence : 33
III.2.4 La majoration de la tarification nationale de référence : 33
III.2.5 Les informations concernant les entreprises assurées : 33
III.2.6 Les informations concernant les prestations demandées : 33
III.2.7 Le taux de remboursement : 34
Partie II : Présentation de la base de données et analyse descriptive 35
Chapitre I : Présentation et analyse de la base de données. 36
I.1. Description de la base de données : 36
I.2 Analyse de la corrélation : 39
Chapitre II: Segmentation et codification des variables 41
II.1. Importance de la segmentation : 42
II.2. Algorithme de CHAID : 42
II.3. Sélection des variables tarifaires : 44
II.4.Segmentation et Codification des variables tarifaires 44
II.4.1 Segmentation et Codification de la variable Taille de groupe : 44
II.4.2 Segmentation et Codification de la variable âge moyen de groupe : 46
II.4.3 Segmentation et Codification de la variable Montant Plafond de remboursement : 48
II.5. Test de Kruskal Wallis: 50
Chapitre III : Description des variables tarifaires 51
III.1. La variable taille du groupe : 51
III.2. La variable âge moyen du groupe: 53
III.3. La variable montant plafond de remboursement : 54
Partie III : Modèles de tarification à priori 56
Chapitre I : Notion d'assurance de groupe. 57
Chapitre II: Présentation de la première approche de tarification à priori. 58
Introduction : 58
II .1 Modélisation de la fréquence : 58
II.1.1 Bref aperçu sur les modèles linéaires généralisés : 58
II-1-2 Application du modèle GLM à la modélisation de la fréquence : 61
II-1-3 Présentation du modèle pour le poste pharmacie : 61

II-1-3-1 Choix du modèle sur la base de la qualité de l'ajustement. 62
II-1-3-2 Test de significativité de Wald : 64
II-1-3-3 Estimation des paramètres du modèle : 64
II.2 Modélisation du coût moyen d'un sinistre : 66
II.2.1 Présentation du modèle pour le poste pharmacie : 66
II.2.1.1 Q-Q plot des lois suspectées 66
II.2.1.2 Récapitulation des modèles de montants de sinistres retenus pour chaque poste: 69
II.3 Application à la détermination de la prime pure : 70
Chapitre III: Présentation de la deuxième approche de tarification à priori. 72
Introduction : 72
III.1 Modélisation de la fréquence : 72
III.1.1 Description du modèle logistique : 72
III.1.2 Estimation des paramètres du modèle : 73
III.1.3 Présentation du modèle pour le poste pharmacie : 75
III.1.3.1 Test de Hosmer et Lemeshow : 76
III.1.3.2 Sélection du modèle : 77
III.1.3.3 Estimation des paramètres du modèle : 78
III.1.3.4 Evaluation du pouvoir discriminant du modèle logistique: 79
III.2 Modélisation de la charge annuelle de sinistralité des polices : 80
III.2.1 Présentation du modèle pour le poste pharmacie : 81
III.2.1.1 Test de Kolmogorov-Smirnov : 82
III.2.2 Récapitulation des modèles retenus de la charge annuelle moyenne de sinistralité de chaque poste : 83
III.3 Application à la détermination de la prime pure : 83
Chapitre IV: Analyse comparative des deux approches de tarification. 86
Partie IV : Application de Tarification des contrats groupe. 88
Chapitre I : Le modèle conceptuel des données : 89
Chapitre II : L'interface principale de l'application : 90
CONCLUSION 92
BIBLIOGRAPHIE : 94
ANNEXES. 96

Liste des abréviations

- **AMO** : Assurance Maladie Obligatoire
- **ANAM** : Agence Nationale de l'Assurance Maladie
- **AUC** : Area Under the Curve
- **CNOPS** : Caisse Nationale des Organismes de Prévoyance Sociales
- **CNSS** : Caisse Nationale du Sécurité Sociale
- **GLM** : Modèles linaires Généralisés
- **IARD** : Incendie, Accident et Risque Divers
- **I S** : Impôt sur les Sociétés
- **MCD** : Modèle Conceptuel de Données
- **ROC** : Receiver Opérating Characteristics
- **SGBD** : Système de Gestion des Bases de Données
- **SQL** : Langage d'interrogation des bases de données
- **T M** : Ticket Modérateur
- **TNR** : Tarification Nationale de Référence

Liste des tableaux

Tableau 2.1 : Informations contenues dans la base de données initiale. 36
Tableau 2.2 les postes de maladie avec leurs codes et les familles d'actes correspondants. 37
Tableau 2.3 : Statistiques descriptives des différents Postes de maladie. 37
Tableau 2.4 Analyse de corrélation entre les variables continues.....41
Tableau 2.6: Les modalités et codes correspondants de la variable taille de groupe. 45
Tableau 2.7 : Les statistiques issues de la segmentation de la variable Taille de groupe 46
Tableau 2.8: Les modalités et codes correspondants de la variable âge moyen de groupe. 47
Tableau 2.9: Les statistiques issues de la segmentation de la variable âge moyen. 48
Tableau 2.10 Les modalités et codes correspondants de la variable montant plafond de remboursement. 49
Tableau 2.11 : Les statistiques issues de la segmentation de la variable montant plafond de remboursement. 49
Tableau 2.12: Résultats du test de kruskal Wallis pour les 3 segmentations effectuées. 50
Tableau 2.3 : Statistiques descriptives des différents Postes de maladie 51
Tableau 2.14 : Répartition des dépenses sanitaires et de fréquence de sinistralité de la variable âge moyen du groupe. 53
Tableau 2.15: Répartition des dépenses sanitaires et de fréquence de sinistralité de la variable Plafond de remboursement. 54
Tableau 3.1 : Tableau récapitulatif de quelques modèles GLM usuels. 59
Tableau 3.2: Analyse de corrélation dans le poste pharmacie. 61
Tableau 3.3: Tableau récapitulatif des modèles pour la fréquence du poste pharmacie. 63
Tableau 3.5 : Estimation des paramètres pour la fréquence du poste pharmacie.....65
Tableau 3.5 (bis) : Estimation des paramètres de l'ajustement du coût moyen en pharmacie.....68
Tableau 3.6 : Tableau récapitulatif de la modélisation des coûts moyens de tous les postes. 69
Tableau 3.7 : Tableau des polices illustrant le calcul de la prime pure.....70
Tableau 3.8 : Application à la détermination de la prime pure selon la première approche de tarification. 71
Tableau 3.9 : Régression backward du poste pharmacie. 75
Tableau 3.10 : Test de Hosmer Lemeshow du poste Pharmacie.....77
Tableau 3.11 : Synthèse de la sélection des variables pour le poste pharmacie. 77
Tableau 3.11 : Estimation des paramètres du modèle logistique pour le poste pharmacie. 78
Tableau 3.12 : Paramètres de l'ajustement normal de la charge annuelle en pharmacie. 82
Tableau 3.13 : Test de Kolmogorov-Smirnov de l'ajustement normal de la charge annuelle. 82
Tableau 3.14: Tableau récapitulatif de la modélisation de la charge de tous les postes. 83
Tableau 3.15 : Application à la détermination de la prime pure pour la deuxième approche.....85
Tableau 3.16: Tableau comparatif des primes obtenues avec les deux(2) approches. 86
Tableau 1.2 : Taux de cotisation de secteur privé. 97
Tableau 1.3 : Taux de couverture de secteur public.....98
Tableau 1.4 : Taux de cotisation de secteur public. 99

Tableau 2.1 (bis) : Actes de Santés et codes correspondants. 100
Tableau 2.5 : Assorti d'analyse de corrélation de Kendall et de Spearman. 101
Tableau 3.17: Analyse de la corrélation de Pearson pour le poste ambulatoire.....103
Tableau 3.18 : Analyse de type III de la variable Taillecode pour le poste ambulatoire. 103
Tableau 3.19 : Estimation des paramètres dans la modélisation de la fréquence des sinistres pour le poste ambulatoire. 104
Tableau 3.20 : Estimation des paramètres dans la modélisation du cout des sinistres pour le poste ambulatoire. 104
Tableau 3.21 : Analyse de la corrélation de Pearson pour le poste pharmacie. 105
Tableau 3.22: Indicateurs d'ajustement du modèle de la fréquence pour le poste pharmacie. 106
Tableau 3.23: Analyse de la corrélation de Pearson pour le poste Hospitalisation. 107
Tableau 3.24 : Indicateurs d'ajustement de la fréquence du poste Hospitalisation. 107
Tableau 3.25 : Analyse de type III de la variable Taillecode du poste Hospitalisation. 108
Tableau 3.26 : Estimation des paramètres du modèle de la fréquence du poste Hospitalisation. 108
Tableau 3.27: Estimation des paramètres dans la modélisation du cout des sinistres pour le poste Hospitalisation. 108
Tableau 3.28: Analyse de la corrélation de Pearson pour le poste Analyses et radios. 109
Tableau 3.29 : Indicateurs d'ajustement de la fréquence du poste Analyses et radios. 110
Tableau 3.30 : Analyse de type III de la variable Taillecode du poste Radios. 110
Tableau 3.31. : Estimation des paramètres du modèle de la fréquence du poste Radios. 111
Tableau 3.32: Estimation des paramètres dans la modélisation du coût des sinistres pour le poste Radios. 111
Tableau 3.33: Analyse de la corrélation de Pearson pour le poste Soins dentaires. 112
Tableau 3.34 : Analyse de type III de la variable Taillecode du poste Soins dentaires. 113
Tableau 3.35 : indicateurs d'ajustement de la fréquence du poste Soins dentaires. 113
Tableau 3.36 : Estimation des paramétres du modèle de la fréquence du poste Soins dentaires. 114
Tableau 3.37: Estimation des paramètres dans la modélisation du cout des sinistres pour le poste Soins dentaires. 114
Tableau 3.38: Analyse de la corrélation de Pearson pour le poste Optique. 115
Tableau 3.39 : indicateurs d'ajustement de la fréquence du poste Optique. 115
Tableau 3.40 : Analyse de type III de la variable Taillecode du poste Optique. 116
Tableau 3.41. : Estimation des paramètres du modèle de la fréquence du poste Optique. 116
Tableau 3.42: Estimation des paramètres dans la modélisation du cout des sinistres pour le poste Optique. 116
Tableau 3.43: Analyse de la corrélation de Pearson pour le poste Autres Prestations. 117
Tableau 3.44 : indicateurs d'ajustement de la fréquence du poste Autres. 118
Tableau 3.45 : Analyse de type III de la variable Taillecode du poste Autres Prestations. 118
Tableau 3.45. : Estimation des paramètres du modèle de la fréquence du poste Autres. 119
Tableau 3.46: Estimation des paramètres dans la modélisation du cout des sinistres pour le poste Autres. 119
Tableau 3.47: Sélection des variables du modèle Logit pour le poste Ambulatoire. 120
Tableau 3.48. : Estimation des paramètres de la régression logistique du poste ambulatoire. 120

Tableau 3.48. : Estimation des paramètres de l'ajustement de la charge annuelle de sinistralité dans le poste ambulatoire. 121
Tableau 3.49: Sélection des variables du modèle Logit pour le poste Hospitalisation. 122
Tableau 3.50 : Estimation des paramètres de la régression logistique du poste Hospitalisation. 122
Tableau 3.51: Estimation des paramètres de l'ajustement de la charge annuelle de sinistralité dans le poste Hospitalisation. 123
Tableau 3.52: Test d'adéquation de la loi Normale à la charge annuelle de sinistralité dans le poste Hospitalisation.....123
Tableau 3.53: Sélection des variables du modèle Logit pour le poste Radios. 124
Tableau 3.54 : Estimation des paramètres de la régression logistique du poste Radios. 125
Tableau 3.55: Estimation des paramètres de l'ajustement de la charge annuelle de sinistralité dans le poste Radios. 126
Tableau 3.56: Test d'adéquation de la loi Normale à la charge annuelle de sinistralité dans le poste Radios.....126
Tableau 3.57: Sélection des variables du modèle Logit pour le poste Soins dentaires. 127
Tableau 3.58 : Estimation des paramètres de la régression logistique du poste Soins dentaires. 127
Tableau 3.59: Estimation des paramètres de l'ajustement de la charge annuelle de sinistralité dans le poste Soins dentaires. 128
Tableau 3.60: Test d'adéquation de la loi Normale à la charge annuelle de sinistralité dans le poste Radios.....128
Tableau 3.61: Sélection des variables du modèle Logit pour le poste Optique. 129
Tableau 3.62 : Estimation des paramètres de la régression logistique du poste Optique. 129
Tableau 3.63: Estimation des paramètres de l'ajustement de la charge annuelle de sinistralité dans le poste Optique.....130
Tableau 3.64: Test d'adéquation de la loi Normale à la charge annuelle de sinistralité dans le poste Optique. 130
Tableau 3.64: Sélection des variables du modèle Logit pour le poste Autres. 131
Tableau 3.65 : Estimation des paramètres de la régression logistique du poste Autres.....131
Tableau 3.66: Estimation des paramètres de l'ajustement de la charge annuelle de sinistralité dans le poste Autres.....132
Tableau 3.67: Test d'adéquation de la loi Normale à la charge annuelle de sinistralité dans le poste Autres. 132

Liste des figures

Figure1.1 : Chiffres d'affaires (en %) des différentes branches du secteur au titre de l'année 2007.20
Figure 2.1 : Répartition des frais de consommation médicale par poste.....38
Figure 2.2 : Répartition des fréquences de maladie par poste.....38
Figure 2.3 : Extrait de la base de données finale (Base1). 40
Figure 2.4 : Arbre de segmentation jusqu'au 3émeniveau pour le ticket modérateur. 44
Figure 2.5 : Arbre de segmentation de la variable Taille de groupe. 45
Figure 2.6 : Arbre de segmentation de la variable âge moyen de groupe. 47
Figure 2.7: Arbre de segmentation de la variable montant plafond de remboursement. 48
Figure 2.9 : Variation des dépenses sanitaires. 52
Figure 2.10: Variation de fréquence de sinistres.....52
Figure 2.11: Variation des dépenses sanitaires. 53
Figure 2.12: Variation de fréquence de sinistres.....54
Figure 2.13 Variation des dépenses sanitaires. 55
Figure 2.14: Variation de fréquence de sinistres.....55
Figure 3.1 : Ajustement Exponentiel des montants de sinistres en pharmacie. 67
Figure 3.2 : Ajustement gamma des montants de sinistres en pharmacie.....67
Figure 3.3 : Ajustement Lognormal des montants de sinistres en pharmacie.....67
Figure 3.4 : Ajustement normal du logarithme des montants de sinistres en pharmacie. 68
Figure 3.6: Courbe ROC du modèle logistique pour le poste pharmacie. 80
Figure 3.7: Q-Q plot de la charge annuelle de sinistres pour le poste pharmacie.....81
Figure 4.1 : le modèle conceptuel des données (MCD). 89
Figure 4.2 : Interface principale de l'application 91
Figure 2.4 : Arbre de segmentation jusqu'au 3éme niveau pour le ticket modérateur a en prenant la fréquence de sinistres comme critère de segmentation. 102
Figure 3.8 : Ajustement Gamma de la transformation logarithmique des montants de sinistre. ..104
Figure 3.9: Ajustement Gamma de la transformation logarithmique des montants de sinistre. ..109
Figure 3.10: Ajustement Gamma de la transformation logarithmique des montants de sinistre. ..112
Figure 3.11: Ajustement Gamma de la transformation logarithmique des montants de sinistre. ..114
Figure 3.12: Ajustement Gamma de la transformation logarithmique des montants de sinistre. ..117
Figure 3.13: Ajustement Gamma de la transformation logarithmique des montants de sinistre. ..119
Figure 3.14: Courbe ROC du modèle Logit pour le poste Ambulatoire. 121
Figure 3.15: Ajustement Normal de la transformation logarithmique de la charge annuelle de sinistralité 121
Figure 3.16: Courbe ROC du modèle Logit pour le poste Hospitalisation. 123

Figure 3.17: Ajustement Normal de la transformation logarithmique de la charge annuelle de sinistralité. 124
Figure 3.18: Courbe ROC du modèle Logit pour le poste Radios. 125
Figure 3.19: Ajustement Normal de la transformation logarithmique de la charge annuelle de sinistralité. 126
Figure 3.20: Courbe ROC du modèle Logit pour le poste Soins dentaires. 127
Figure 3.21: Ajustement Normal de la transformation logarithmique de la charge annuelle de sinistralité. 128
Figure 3.22: Courbe ROC du modèle Logit pour le poste Optique. 129
Figure 3.23: Ajustement Normal de la transformation logarithmique de la charge annuelle de sinistralité. 130
Figure 3.24: Courbe ROC du modèle Logit pour le poste Autres.....131
Figure 3.25: Ajustement Normal de la transformation logarithmique de la charge annuelle de sinistralité. 132

Introduction

Le marché marocain des assurances est le deuxième en Afrique. Ce marché vit depuis des années au rythme de mouvement de concentration, de fusion, de prise de participation et de processus de libéralisation, une libéralisation qui promet d'être complète dans l'année 2010, date d'entrée en vigueur de l'accord entre le Maroc et l'union Européenne, et qui vise le démantèlement des droits de douanes et une libéralisation totale du marché marocain.

La branche assurance vie constitue l'un des principaux piliers de ce secteur, la concurrence sur le marché entre les compagnies d'assurance devra y être ardue. Ainsi, pour garder ses parts de marché et pouvoir attirer aussi une bonne partie de nouveaux assurés, chaque compagnie devra diversifier ses prestations, notamment en assurance santé. Surtout après l'entrée en vigueur de l'assurance maladie obligatoire (AMO) qui va s'étendre à une large population marocaine et qui bénéficiera des prestations en assurance de base en payant des primes assez modestes, d'où l'obligation pour ces compagnies de développer une nouvelle prestation en assurance santé dite assurance maladie complémentaire.

La Mutuelle Marocaine d'Assurance Agricole et la Mutuelle Centrale Marocaine d'Assurance, dernièrement fusionnées sous la forme de société de personnes sous la dénomination : la MAMDA-MCMA, fait partie de ce secteur. Elle devra être en mesure de pratiquer un tarif adéquat à ce nouveau type de prestation telle que l'assurance maladie complémentaire. Ce tarif sera le résultat d'application des méthodes de calculs actuariels et des méthodes de statistiques de plus en plus avancées qui estiment les montants de primes le plus précisément possible et différencie celles-ci en fonction des caractéristiques des assurés, car il n'est désormais plus concevable qu'une compagnie d'assurance se passe de ces techniques dans l'environnement concurrentiel et déréglé qui voit le jour.

Ainsi, en vue de pratiquer les méthodes actuarielles et statistiques précitées, vient ce travail qui comporte quatre parties :

La première partie est consacrée aux généralités et à la présentation de secteur des assurances au Maroc, notamment en assurance maladie obligatoire, et en assurance maladie complémentaire.

La deuxième partie présente la base de données issue du portefeuille de la MAMDA-MCMA et mène une description de ses différentes composantes ainsi qu'une segmentation de

ce portefeuille et des variables tarifaires les plus pertinentes. Cette segmentation consiste à l'affectation de chaque assuré à la classe de risque correspondante en fonction des caractéristiques observables et spécifiques au risque assuré.

Dans la troisième partie, on effectuera la tarification proprement dite des contrats groupe selon les deux approches dont l'une consiste à modéliser la fréquence de sinistres par une loi binomiale négative, et l'autre par une régression logistique.

On clôturera notre travail par une présentation sommaire d'une application informatique qui a été conçue dans le but d'automatiser le fonctionnement des modèles élaborés et retenus par notre travail de recherche, en utilisant un logiciel SGBD et plus précisément en implémentant le modèle conceptuel de données (MCD) avec les outils du logiciel ACCESS, ce qui fera l'objet de la quatrième et la dernière partie de ce Travail.

Partie I

Généralités & Présentation

Cette première partie se compose de trois Chapitres :

Le Chapitre I qui traitera les composantes du secteur marocain des assurances, ainsi que les nouveautés qu'avait connu ce marché durant les deux dernières années. Le Chapitre II dans lequel on présentera l'organisme d'accueil, puis le Chapitre III qui donnera successivement des aperçus sur l'Assurance Maladie Obligatoire de base (AMO), et sur l'Assurance Maladie Complémentaire.

Chapitre I : Les caractéristiques du secteur des assurances au Maroc

I .1 : les composantes du secteur marocain des assurances

Le secteur marocain des assurances compte 17 compagnies d'assurance à savoir : *Atlanta, Axa Assistance Maroc, Axa Assurance Maroc, CAT, CNIA Assurance, Essaada, Euler Hermès ACMAR, ISAAF Mondiale Assistance, MAMDA- MCMA, Maroc Assistance Internationale, Marocaine Vie, MATU, RMA Watanya, Sanad, Wafa Assurance et Zurich*.

Le marché marocain des assurances est le deuxième en Afrique, avec un chiffre d'affaires qui est passé de 16,7 milliards de DH en 2007 à environ 18,9 milliards de DH en 2008, soit une hausse de 13%. L'assurance non-vie est toujours prédominante dans le portefeuille du secteur avec 13 milliards de DH, soit une part de marché de plus de 68%. Actuellement, c'est Wafa assurance qui arrive en premier avec 4,1 milliards de DH, détrônant ainsi RMA Watanya , jusque-là leader et qui suit à 4 milliards de DH, alors que AXA assurance Maroc arrive en troisième position avec 2,5 milliards de DH.

I.2 Les apports et les nouveautés des années 2007 et 2008

Durant les années 2007 et 2008, le secteur des assurances a connu plusieurs événements marquants :

- En 2007, l'introduction de 17% du capital d'ATLANTA en bourse, troisième compagnie d'assurance qui ouvre son capital au public après la Marocaine Vie et Wafa Assurance.
- Le deuxième événement marquant de l'année 2007 est la prise de participation par la Banque Centrale Populaire de 43,54% du capital de la Marocaine Vie. Cette prise de participation s'est faite par cession de bloc avant d'être annulée durant le premier trimestre de l'année 2008.
- En 2008 rachat d'Essaâda par le groupe Saham, actionnaire de référence de CNIA.
- En 2008 aussi, la loi de finances a introduit des nouvelles dispositions fiscales qui intéressent le secteur des assurances ; il s'agit de :

- La revalorisation du taux de taxe sur les contrats d'assurances (3,50% pour la capitalisation au lieu de 3,45% ; 7% pour les facultés maritimes au lieu de 6,90% ; 14% pour les autres assurances IARD au lieu de 13,80%).
- La suppression de la déductibilité de l'impôt sur le revenu des versements exceptionnels sur salaire au titre des contrats d'assurance retraite.
- La réduction du taux de l'IS qui passe de 39,6% à 37% dès l'année 2008, avec l'engagement du gouvernement de le porter à 35% à compter de 2009.

I .3 Chiffres d'affaires des différentes branches du secteur marocain des assurances au titre de l'exercice 2007 :

Branche	chiffre d'Affaire(DH)	Contribution
Automobile	5 427,90	30,70%
Assurances Individuelles et Capitalisation	4 238,40	24,00%
Accidents Corporels	2 228,20	12,60%
Accidents du Travail	1 476,90	8,40%
Assurances Groupes et Capitalisation	1 351,80	7,70%
Incendie	799,8	4,50%
Transport	682,2	3,90%
Assistance - Crédit - Caution	453,7	2,60%
Responsabilité Civile Générale	342,8	1,90%
Contrats à Capital Variable	273,7	1,50%
Assurances des Risques Techniques	235	1,30%
Autres Opérations Non Vie	124,2	0,70%
Acceptations en réassurance	32,5	0,20%
TOTAL	**17 667,10**	**100,00%**

Tableau1.1 : Chiffres d'affaires du secteur marocain des assurances durant l'année 2007.

Source : Fédération Marocaine des Sociétés d'Assurances et de Réassurance.

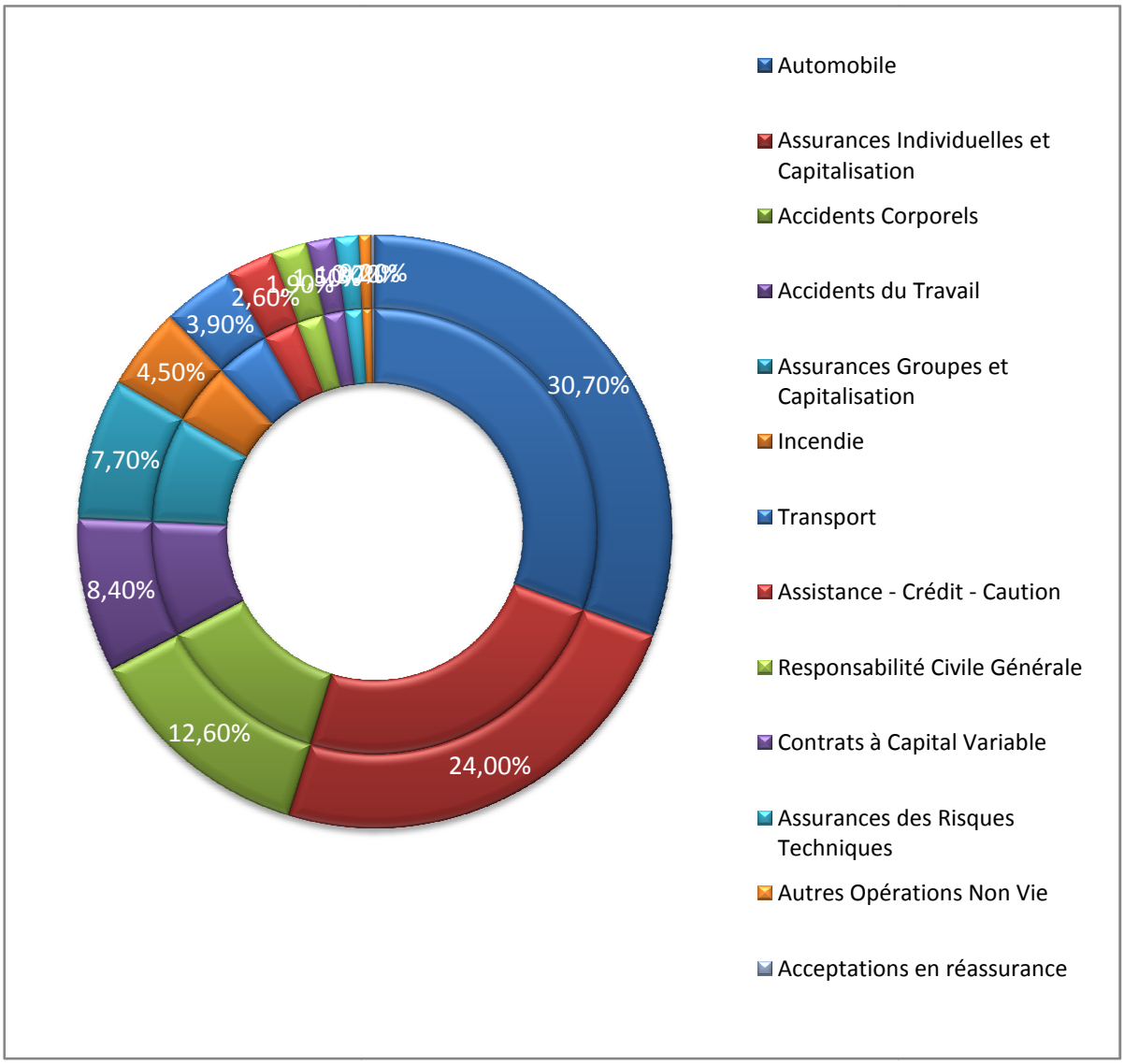

Figure1.1 : Chiffres d'affaires (en %) des différentes branches du secteur au titre de l'année 2007.

C'est l'assurance automobile qui arrive en premier rang avec 30,70% du total du chiffre d'affaires de tout le secteur, suivi de l'assurance individuelle et capitalisation avec 24%, et l'assurance des accidents corporels avec 12,60%. Quant aux pourcentages des autres branches, ils ne dépassent pas un seul chiffre.

Une des composantes la plus connue de ce secteur est la compagnie MAMDA-MCMA dont la représentation fait l'objet du chapitre suivant.

Chapitre II : Présentation de l'organisme d'accueil MAMDA-MCMA

II.1 Fiche technique de la MAMDA-MCMA

- Forme juridique	: Société de personnes
- Objet social	: mutuelle d'assurance des biens et des personnes
- Adresse	: 16, Rue Abou Inane, Rabat
- Actionnaires	: ……
- Capital social	: ……
- Part de marché	: ……
- Chiffres d'affaires	: 890 millions de dirhams.
- Fonds propres	: ……
- Effectifs	: 300 personnes
- Réseaux de distribution	: 30 Bureaux régionaux tous au Maroc.

II.2 Activités et performance de la MAMDA-MCMA

II.2.1 Activités de la MAMDA-MCMA

Les Assurances Mutuelles MAMDA et MCMA sont l'émanation d'un passé mutualiste datant des années vingt (Dahir du 30 Octobre 1920) et a acquis ses assises juridiques par la loi 17-99 portant code des assurances venue consacrer le caractère « d'entreprise d'assurances » de la société mutuelle à l'instar de la société anonyme. Elle opère sans restriction dans tous les domaines d'assurance et est soumise aux mêmes obligations et contrôles.

La MAMDA est une entreprise d'assurance immatriculée au registre de commerce dotée d'un fonds d'établissement de 100 millions DHS et dont les statuts particuliers permettent de distribuer à ses assurés soit une ristourne de mutualité prise sur les excédents de l'exercice soit de procéder à une réduction de la cotisation et de bénéficier ainsi d'un tarif préférentiel.

La MAMDA existe sous sa forme actuelle depuis 1962. Sa vocation lui confère un domaine d'intervention concentré sur l'agriculture. Elle couvre ainsi les risques liés à l'ensemble des activités agricoles.

La MCMA, qui est une émanation de la MAMDA, quant à elle, jouit d'un champ d'action élargi et couvre l'ensemble des risques classiques liés à tout type d'activité d'assurances en dehors de l'agriculture.

La vocation de la MCMA est donc généraliste et s'adresse à tout le monde : groupements (Mutuelle d'Assurances des Enseignements du Maroc dite MAEM, la Garantie des Enseignants du Supérieur dite GASUP, Mutuelle d'Assurances des Transports au Maroc dite MATRAM, etc.…), entreprises publiques ou privées et par extension toute personne physique désirant s'assurer.

Aussi, la recherche des groupements a été toujours de mise à l'instar de l'APEFEL (association des producteurs exportateurs des fruits et légumes), le partenariat avec SOCIALAGRI pour le personnel relevant de l'agriculture et la convention avec l'Ordre National des Vétérinaires et du Club Méditerranée.

De même, la MCMA fait actuellement de la bancassurance avec des partenaires solides sur le marché tels que le Crédit Agricole du Maroc, la Banque Centrale Populaire et Attijari Wafabank.

La MAMDA et la MCMA, sont des mutuelles, organisées juridiquement sous forme de sociétés de personnes et non de capitaux. Elles n'ont donc aucun but lucratif. Leur objectif majeur est d'offrir aux assurés le meilleur service au meilleur prix.

Les Assurances Mutuelles MAMDA et MCMA occupent une position confortable et originale au sein du marché marocain de l'Assurance.

La MAMDA est le premier assureur agricole du marché ; elle est consciente du rôle qu'elle joue dans le développement de l'agriculture en général ; elle gère pour le compte de l'Etat et ce depuis 1994 le programme de garantie des céréales contre les effets de la sécheresse, à titre bénévole.

La MCMA quant à elle, se différencie par une politique généraliste, toutes cibles confondues.

Aujourd'hui, MAMDA et MCMA regroupent près de 160.000 Sociétaires et emploient plus de 300 personnes.

L'activité financière du groupe en termes de placements mobiliers et immobiliers fait de lui l'un des meilleurs opérateurs du marché financier marocain. La MAMDA et La MCMA figurent aujourd'hui parmi les grands opérateurs boursiers de la place de Casablanca. Le chiffre d'affaires du groupe est de près de 890 millions de dirhams. Les activités financières de la MAMDA et de la MCMA se concentrent sur la gestion d'un important portefeuille de participations auprès de nombreuses sociétés (de crédits et de financement, immobilier, agro–industrie). Les qualités de gestion financière des assurances MAMDA – MCMA, reconnues par l'ensemble du marché financier marocain, dotent ces deux mutuelles d'une personnalité originale marquée par la rigueur, l'expertise et l'audace.

II.2.2 Expériences de la MAMDA-MCMA

La Mutuelle Centrale Marocaine d'Assurances jouit d'une longue expérience dans le domaine de la Prévoyance, la gestion de produits d'épargne et de la capitalisation depuis les années 70.

La M.C.M.A opère depuis 40 ans avec des partenaires solides, et offre des prestations d'assurance vie de dernière génération; il s'agit de :

- Banque Commerciale du Maroc (BCM)
- Attijari Wafabank
- Crédit Agricole du Maroc
- Club Méditerranée
- Offices Régionaux de Mise en Valeur Agricole
- Groupement du supérieur « GASUP »
- Union Nationale du Préscolaire (UNAPS)
- Association des Producteurs et Exportateurs des Fruits et Légumes(APEFEL).

II.2.3 Les Réassureurs de la MAMDA-MCMA

Le Groupe MAMDA – MCMA traite avec des réassureurs d'une réputation internationale issue, à la fois de leur solidité financière et leur expertise largement reconnue comme :

Société Centrale de Réassurance (S.C.R.)	MAROC
Mutuelle Centrale de Réassurance (M.C.R.)	FRANCE
SCOR	FRANCE
PARIS RE	FRANCE
SWISS RE	SUISSE
HANNOVER RE	ALLEMAGNE
BEST RE	TUNISIE

Le groupe MAMDA-MCMA s'intéresse actuellement plus que jamais, comme d'ailleurs le cas de la plus part des compagnies marocaine d'assurance, à l'assurance santé surtout avec l'instauration d'une assurance maladie Obligatoire et d'une Assurance maladie complémentaire. Le chapitre suivant traitera avec plus de détails ces deux types d'assurance.

Chapitre III : Assurance maladie Obligatoire & Assurance maladie complémentaire.

III-1 Assurance maladie Obligatoire

III-1-1 Présentation :

Afin de concrétiser l'engagement de l'Etat, qui consacre le principe du droit à la santé tel que prévu par les conventions internationales la loi 65-00 portant code de la couverture médicale de base constitue le parachèvement de l'expérience du Maroc en matière de couverture médicale et consolide les droits acquis par les citoyens marocains bénéficiant

d'une assurance maladie. Cette assurance sera progressivement étendue à l'ensemble des citoyens, toutes catégories sociales confondues.

A cette fin, un système obligatoire de couverture médicale de base a été mis en place en vue d'atteindre l'accès universel aux soins.

Dans ce cadre la loi 65-00 portant code de la couverture médicale de base institue une assurance maladie obligatoire de base (AMO), qui est fondée sur le principe contributif et la mutualisation des risques.

L'article 1er du décret 2-05-738 fixant les conditions d'affiliation et d'immatriculation au régime d'Assurance Maladie Obligatoire de base précise que, sous réserve des dispositions de l'article 114 de la loi 65-00, sont considérés comme employeurs affiliés d'office au titre de l'AMO :

- Les employeurs affiliés à la CNSS au titre des prestations de sécurité sociale,
- Les administrations de l'Etat, les collectivités locales, les établissements publiques et les personnes morales de droit public dont les fonctionnaires et agents sont à la date de publication du décret ; soit le 18 août 2005, immatriculés à la CNOPS, les organismes gérant les régimes de pensions.

III.1.2 Les prestations fournis par l'AMO :

L'assurance maladie obligatoire régie par la CNSS couvre quatre catégories de prestations :

- Les maladies graves ou invalidantes nécessitant des soins de longue durée ou particulièrement couteux,
- Les prestations destinées à l'enfant dont l'âge est inférieur ou égale à 12 ans,
- Le suivi de la grossesse, de l'accouchement et de ses suites,
- L'hospitalisation.

➢ ***Les maladies graves ou invalidantes nécessitant des soins de longue durée ou particulièrement couteux :***

- Actes de médecine générale, de spécialités médicales et chirurgicales et Analyse de biologie médicale ;
- Radiologie et imagerie médicale ;

- Explorations fonctionnelles ;
- Hospitalisation ;
- Médicaments admis au remboursement ;
- Sang et ses dérivés labiles ;
- Soins bucco-dentaires ;
- Dispositifs médicaux et implants nécessaires aux actes médicaux et chirurgicaux admis au remboursement ;
- Acte de rééducation fonctionnelle et de kinésithérapie ;
- Actes paramédicaux ;
- Appareils de prothèse et d'orthèse médicales admis au remboursement ;
- Lunetterie médicale.

Les prestations ci-dessus, sont couvertes par l'AMO de base quelles soient dispensées à titre ambulatoire ou dans le cadre de l'hospitalisation.

- ***Les prestations destinées à l'enfant dont l'âge est inférieur ou égal à 12 ans :***

  - Soins préventifs et curatifs liés aux programmes prioritaires entrant dans le cadre de la politique sanitaire de l'Etat ;
  - Actes de médecine générale et de spécialités médicales et chirurgicales ;
  - Soins liés à l'hospitalisation et aux interventions chirurgicales, y compris les actes de chirurgie réparatrice ;
  - Analyse de biologie médicale ;
  - Radiologie et imagerie médicale ;
  - Explorations fonctionnelles ;
  - Médicaments admis au remboursement ;
  - Poche de sang humain et dérivés sanguins ;

- Dispositifs médicaux et implants nécessaires aux actes médicaux et chirurgicaux compte tenu de la nature de la maladie ou de l'accident et du type de dispositifs ou d'implants ;
- Appareils de prothèse et d'orthèse médicales admis au remboursement ;
- Lunetterie médicale ;
- Soins bucco-dentaires ;
- Orthodontie pour les enfants ;
- Acte de rééducation fonctionnelle et de kinésithérapie ;
- Actes paramédicaux.

➢ ***Le suivi de la grossesse, de l'accouchement et de ses suites :***

- L'accouchement et ses suites ;
- Les actes médicaux et chirurgicaux tels qu'ils sont définis à la nomenclature des actes professionnels et à la nomenclature des actes de biologie médicale ;
- Les Médicaments admis au remboursement ;
- Le Sang et ses dérivés labiles ;
- Les actes paramédicaux ;
- L'acte de rééducation fonctionnelle et de kinésithérapie.

➢ ***Hospitalisation :***

L'ensemble des prestations et soins rendus dans ce cadre, y compris les actes de chirurgie réparatrice.

III.2.1.1 Prestations couvertes et taux de Cotisation du secteur privé CNSS :

Selon l'article premier du décret n° 2-05-737, les groupes de prestations couvertes par la Caisse Nationale de Sécurité Sociale au titre de l'assurance maladie obligatoire de base sont définis comme suit :

✓ En ce qui concerne les maladies graves ou invalidantes nécessitant des soins de longue durée ou particulièrement couteux, les prestations en médicaments requises sont les suivantes :

- Actes de médecine générale et de spécialités médicales et chirurgicales, Analyse de biologie médicale ;
- Radiologie et imagerie médicale ;
- Explorations fonctionnelles ;
- Hospitalisation ;
- Médicaments admis au remboursement ;
- Sang et ses dérivés labiles ;
- Soins bucco-dentaires ;
- Dispositifs médicaux et implants nécessaires aux actes médicaux et chirurgicaux admis au remboursement ;
- Acte de rééducation fonctionnelle et de kinésithérapie ;
- Actes paramédicaux ;
- Appareils de prothèse et d'orthèse médicales admis au remboursement ;
- Lunetterie médicale.

Les prestations ci-dessus sont couvertes qu'elles soient dispensées à titre ambulatoire ou dans le cas d'hospitalisation.

- ✓ En ce qui concerne l'enfant dont l'âge est inférieur ou égal à 12ans :

L'ensemble des prestations est définis à l'article 7 de la loi 65-00. A ce titre, il convient de signaler la parution de l'arrêté du ministère de la santé n°2563-05, du 5 septembre 2005, qui édicte les mesures nécessaires au suivi médical de l'enfant dans le cadre de l'AMO.

- ✓ En ce qui concerne le suivi de la grossesse, l'accouchement et ses suites :

Les actes médicaux et chirurgicaux tels qu'ils sont définis à la nomenclature des actes de biologie médicale ainsi que les médicaments admis au remboursement; le sang et ses dérivées labiles ; les actes paramédicaux et ; le cas échéant, les actes de rééducation fonctionnelle et de kinésithérapie. Il convient de signaler à ce niveau la parution au bulletin officiel n°5384, du 5 janvier 2006, du l'Arrêté et du ministre de la santé n° 2509-05 qui fixe les conditions et les épisodes de suivi médical de la grossesse, de l'accouchement et de ses suites.

- ✓ En ce qui concerne les hospitalisations, l'ensemble des prestations et soins entrant dans ce cadre, y compris les actes de chirurgie réparatrice.

2-1-1-1 Taux de cotisation de secteur privé :

Les différents taux de cotisation de ce secteur sont représentés dans l'annexe 1.

2.1.1.2 Taux de couverture de secteur privé :

Selon l'article 2 et l'article 3 du décret n° 2-05-737 du 11 joumada II 1426(18 juillet 2005) :

- Le taux de couverture des prestations servis par la CNSS est fixé à 70% de la tarification nationale de référence, établie conformément à l'article 12 de la loi 65-00.
- Ce taux est porté à 90% pour les maladies graves ou invalidantes nécessitant des soins de longue durée ou particulièrement couteux, voir arrêté n°2518-05, lorsque les prestations y afférentes sont dispensées dans les hôpitaux publics, les établissements publics de santé et de services sanitaires relevant de l'Etat.
- La prise en charge des maladies graves ou invalidantes nécessitant des soins de longue durée ou particulièrement coûteux englobe l'ensemble des prestations médicalement requises par l'état de santé du bénéficiaire, y compris la transplantation d'organes et de tissus.

III.2.1.2 Prestations couvertes et taux de couverture du secteur public CNOPS

III.2.1.2 .1 Taux de couverture de secteur public :

Les différents taux de couverture de ce secteur sont représentés dans l'annexe 2.

En cas de maladie grave ou invalidante nécessitant des soins de longue durée ou partiellement payés par l'assuré selon le type de maladies telles que prévues dans la liste arrêtée par le ministère de la santé (l'article 2 du décret n° 2- 05-736).

La part restant à la charge de l'assuré ne peut être supérieure à 10% de la tarification nationale de référence pour ces maladies.

Toutefois, l'exonération de la part restant à la charge de l'assuré est totale pour les soins onéreux (l'article 2 du décret n° 2- 05-736).III.2.1.2 .2 Taux de cotisation de secteur public :

Les différents taux de cotisation de ce secteur sont représentés dans l'annexe 3.

III.1.3 Les bénéficiaires :

III.1.3.1 Secteur public :

Sont soumis au régime AMO dans le secteur public :

Les fonctionnaires, les agents temporaires, occasionnels de l'Etat, les magistrats, le personnel d'encadrement et de rang des Forces auxiliaires, le corps des administrateurs du ministère de l'intérieur, ainsi que le personnel des collectivités locales, des établissements publics, les personnes morales de droit public; les catégories des agents journaliers qui ont été fixées par décret n°2-05-739.

Les bénéficiaires de ce régime s'étendent notamment à leurs ayants droit :

- Le(s) conjoint(s) de l'assuré ;(le conjoint est assuré seulement s'il ne bénéficie d'aucun autre régime).
- Les enfants à la charge de l'assuré, âgés de 21 ans au plus; cette limite d'âge est prorogée jusqu'aux 26 ans pour les enfants non mariés poursuivant des études supérieures, à condition d'en apporter la justification.
- Les enfants pris en charge conformément à la législation en vigueur.
- Sont considérés comme personnes à charge sans limite d'âge, les enfants de l'assuré atteints d'un handicap physique ou mental et les enfants pris en charge et qui sont dans l'impossibilité totale, permanente et définitive de se livrer à une activité rémunérée (Article 5 de la loi 65-00).
- L'assuré pourrait demander l'extension du bénéfice du régime d'assurance maladie obligatoire de base dont il relève à son père et à sa mère, à condition de prendre en charge la cotisation les concernant dés que le décret d'application sera publié à cet effet.
- Lorsque les parents sont l'un et l'autre assurés en vertu des dispositions de la présente loi, les enfants sont déclarés à l'organisme assureur du père.
- En cas de dissolution de mariage, la déclaration doit être faite auprès de l'organisme assureur de l'ex-conjoint auquel la garde des enfants est confiée. Si la garde des enfants est confiée à une personne autre que la mère ou le père, les enfants conservent le bénéfice du régime d'assurance maladie obligatoire de base de l'un des parents assurés.

- Lorsque seul l'un des parents est assuré et en cas de dissolution de mariage, la déclaration doit être faite auprès de l'organisme assureur de l'ex-conjoint. (Article 6 de la loi 65-00).
- Les personnes titulaires de pensions de retraite, de vieillesse, d'invalidité ou d'ayant cause allouées notamment en vertu des régimes suivants :

✓ Le régime des pensions civiles institué par la loi n° 011-71, telle qu'elle a été modifiée et complétée, en ce qui concerne le personnel civil de l'Etat et des collectivités locales ;

✓ Le régime des pensions militaires institué par la loi n° 013-71,telle qu'elle a été modifiée et complétée, étendue par le dahir portant loi n° 1-71-92 en ce qui concerne les personnels d'encadrement et de rang des Forces auxiliaires ;

✓ Le régime des collectivités locales de retraite institué par le dahir portant loi n°1-77-216, telle qu'elle a été modifiée et complétée .

III-1-3-2 Secteur privé:

Dans ce secteur, les bénéficiaires du régime AMO sont les mêmes que ceux indiqués dans les huit premiers points du secteur public.

Il faut signaler qu'à titre transitoire, et pendant une période de cinq ans renouvelables , à compter de la date de publication des décrets réglementaires de la loi 65-00, soit le 18 août 2005, les organismes publiques 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 assurés 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 conformément aux dispositions de l'article 2 du décret 2-05-738.

- 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 d'une couverture (article 144 de la loi 65-00).

- Dans tous les cas, sont conservés les droits acquis par les bénéficiaires des dits régimes, tant pour la partie prenante de charge les cotisations que pour le taux de couverture dont ils bénéficient (article 144 de la loi 65-00).

La mise en œuvre progressive de l'AMO fait que l'entrée en vigueur s'applique dans un premier temps aux salariés des secteurs public et privé, à leurs ayants droit ainsi qu'aux pensionnés des deux secteurs public et privé.

D'autres régimes seront mis en place pour étendre le bénéfice de l'AMO à d'autres catégories sociales, il s'agit notamment des :

- Travailleurs indépendants,
- Personnes exerçant une profession libérale,
- Personnes exerçant une activité non salariée,
- Anciens résistants et membres de l'armée de libération.
- Etudiants dans la mesure où ils n'en bénéficient pas en vertu d'un autre titre.

Dans la majorité des cas, l'AMO ne rembourse qu'une partie des frais médicaux engagés par l'assuré. Le reliquat non remboursé de ces frais, pourra faire l'objet d'une assurance maladie complémentaire.

III.2 Assurance maladie complémentaire

III.2.1 Qu'est ce que l'assurance maladie complémentaire ?

Une assurance maladie complémentaire ou tout simplement une complémentaire santé a pour objet de prendre en charge tout ou partie des dépenses de santé qui ne sont pas couvertes par l'AMO.

Généralement, les soins, les séjours hospitaliers et les visites chez les médecins ou chez les pharmaciens coûtent cher. Souscrire une complémentaire santé permet donc aux assurés d'être mieux remboursé de leurs dépenses de santé. La CNSS, la CNOPS ainsi que d'autres assureurs ne remboursent pas totalement ces différents frais; une complémentaire santé permet donc de compenser la différence entre le coût d'une prestation de santé et son remboursement par ces différentes caisses.

III.2.2 Le remboursement du Ticket Modérateur :

Le Ticket Modérateur est le reliquat des frais supportés par le malade après le remboursement du régime de base. Les compagnies d'assurances s'engagent de rembourser

un pourcentage contractuel de ce ticket modérateur jusqu'à hauteur de la tarification nationale de référence.

III.2.3 La Tarification Nationale de Référence :

La Tarification Nationale de Référence est fixée conformément à l'article 12 de la loi n° 65-00 portant code de la couverture médicale de base. Il s'agit de grilles tarifaires fixant le niveau des remboursements des différents actes médicaux.

III.2.4 La majoration de la tarification nationale de référence :

Certaines entreprises estimant que la tarification nationale de référence est en deçà des tarifs réels des différents actes médicaux se voient proposer par les compagnies d'assurances des coefficients multiplicateurs en fonction des options choisies. Ces coefficients permettent de se rapprocher des frais réels exposés facturés par le médecin traitant, ou suite à un séjour en clinique, qui peuvent dépasser la tarification nationale de référence.

Plusieurs coefficients multiplicateurs sont proposés en contrepartie d'une majoration de primes. Il faut bien noter que quelque soit le taux multiplicateur choisi, la complémentaire santé ne remboursera jamais au delà des frais engagés.

III.2.5 Les informations concernant les entreprises assurées :

Pour établir une tarification adaptée à la situation de chaque entreprise assurée, l'assureur a besoin de connaitre :

- La taille de chaque groupe d'assurés, notamment le nombre de collaborateurs ; le nombre de conjoints ; le nombre d'enfants (avec précision de ceux âgés de plus de 12ans);
- l'âge moyen de chaque groupe d'assurés.

III.2.6 Les informations concernant les prestations demandées :

Il s'agit de :

- Du taux de remboursement ;
- Du plafond général de remboursement ;
- Du plafond de chaque rubrique (Optique, Soins dentaires, hospitalisation, pharmacie,…etc.)

III.2.7 Le taux de remboursement :

Le taux de remboursement est toujours lié à la tarification nationale de référence, il est de :

- 70% de la tarification nationale de référence pour tous les actes non liés aux maladies graves ou invalidantes nécessitant des soins de longue durée ou particulièrement couteux ;

- 90% de la tarification nationale de référence pour tous les actes non liés aux maladies graves ou invalidantes nécessitant des soins de longue durée ou particulièrement couteux ; lorsque les prestations y afférentes sont dispensées dans les hôpitaux publics, les établissements publics de santé et les services sanitaires relevant de l'Etat.
- La prise en charge des maladies graves ou invalidantes nécessitant des soins de longue durée ou particulièrement couteux englobe l'ensemble des prestations médicalement requises par l'état de santé du bénéficiaire y compris la transplantation d'organes et de tissus.

Partie II

Présentation de la base de données et analyse descriptive.

Dans cette deuxième partie, on présentera la base de données qui a servi comme assise dans tout notre travail de recherche, avant de procéder à l'analyse descriptive de ses différentes composantes (chapitre I), puis on procédera aux segmentations et aux codifications des variables tarifaires les plus pertinentes (chapitre II), et aux descriptions plus détaillées de ses différentes variables (chapitre III).

Chapitre I : Présentation et analyse de la base de données.

I.1. Description de la base de données :

L'étude porte sur un portefeuille risque maladie contenant les prestations par acte médical pour une période de quatre ans allant de l'année 2004 jusqu'à 2007 de la compagnie MAMDA-MCMA.

Le fichier initial, que nous avons reçu de la part de la compagnie MAMDA-MCMA comporte dans son état brut environ 104002 observations correspondant aux déclarations de sinistres des assurés et leurs indemnisations. Chaque observation contient les 15 informations suivantes (Voir tableau 2.1) :

Libellé d'information	Etiquette
L'année de survenance de la prestation	ANSURV
Le code de l'entreprise adhérant	CDEEST
Le code de bureau d'inscription	CDEBUR
Le numéro de Police (ou de contrat)	NUMPOL
Le numéro d'adhérant	NUMADH
Le numéro de bénéficiaire	NUMBEN
Le code acte de maladie	CDEACT
Le libellé d'acte de maladie	LIBACT
Le Prix unitaire	PRIXUNT
Les frais réels	FRAISREL
La base de remboursement	BASREM
Le coefficient d'acte	COEACT
Le montant remboursé	MNTREM
Le montant plafond de remboursement	MNTPLF
La taille de groupe (de l'entreprise)	TAILLE

Tableau 2.1 : Informations contenues dans la base de données initiale.

Source : Base de données de la MAMDA-MCMA.

Cette base de données contient **63 actes de maladie** (ou actes de santé).Chaque acte portera un code. Ces différents actes de maladies ainsi que les codes correspondant sont représentés dans le tableau 2.1 (bis) en annexe 1.

Ensuite, il a fallu pour affiner l'étude, regrouper les 63 actes de maladie en **sept familles d'actes ou postes**. Il faut noter que le regroupement des actes s'est opéré en tenant compte de la nature de chaque acte et aussi et des considérations commerciales de ces types de produits.

Le tableau 2.2 montre les postes de maladie ainsi constitués avec leurs codes et les familles d'actes correspondants.

[redacted]

Tableau 2.2 les postes de maladie avec leurs codes et les familles d'actes correspondants.

L'analyse descriptive du portefeuille de la compagnie, agrégé selon les sept postes de maladie, a pu mettre en évidence des répartitions inégales en matière de frais de consommation médicale, relatifs à chaque poste, réellement engagés par les assurés et en matière de fréquences de sinistralité aussi.

Le tableau 2.3 ainsi que les figures 2.1 et 2.2 montrent bien ces répartitions :

Poste	Somme des Frais réels en (DH)	Somme des Frais réels en %	Fréquences	Fréquences en %
1 : Ambulatoire	6277676,44	13%	35095	34%
2 : Pharmacie	15476391,74	33%	41287	40%
3 : Hospitalisation	5381521,23	11%	3398	3%
4 : Analyse & Radios	5727896	12%	12555	12%
5 : Soins & Prothèses	8935592,23	19%	6598	6%
6 : Optique	4360827,69	9%	4527	4%
7 : Autres	1260778,79	3%	434	0%
Total	**47420684,12**	**100%**	**103894**	**100%**

Tableau 2.3 : Statistiques descriptives des différents Postes de maladie.

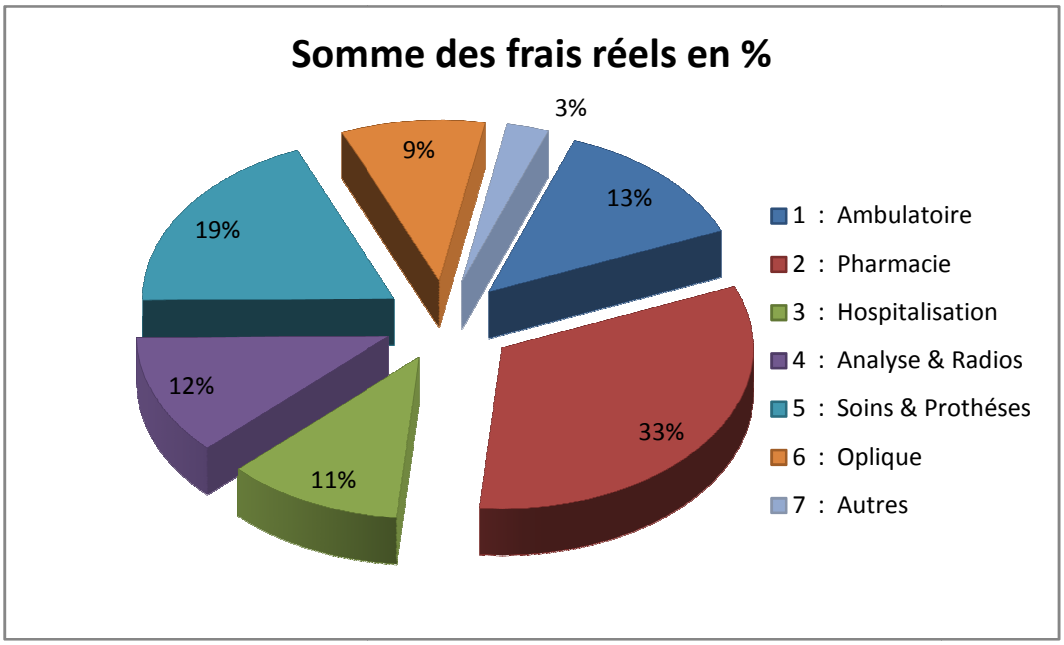

Figure 2.1 : Répartition des frais de consommation médicale par poste.

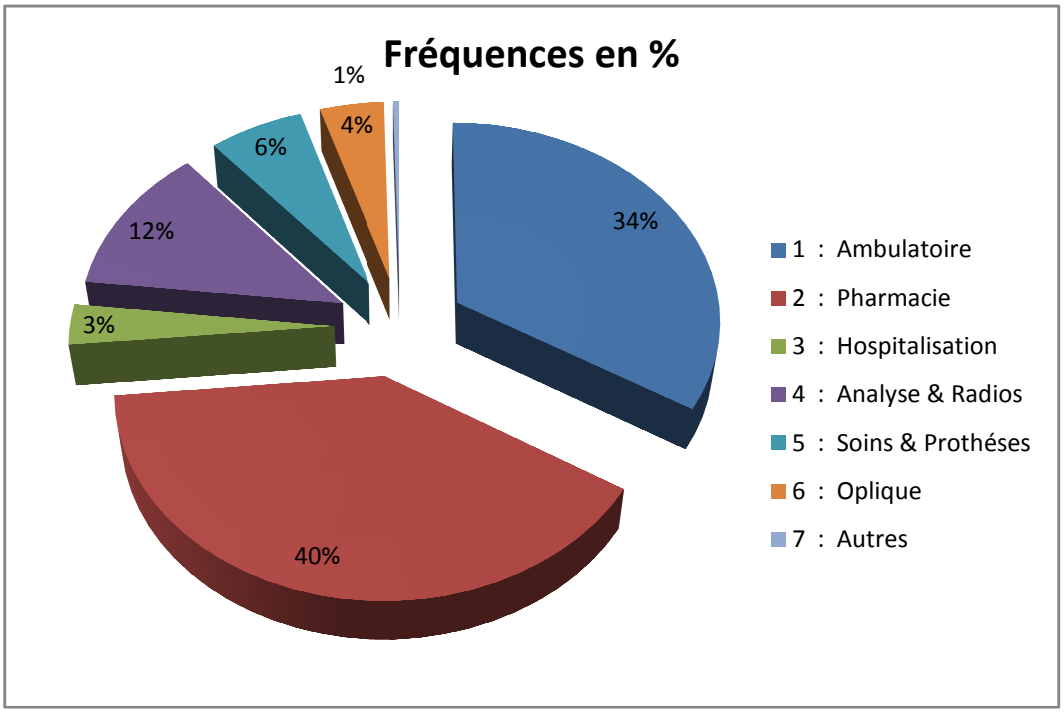

Figure 2.2 : Répartition des fréquences de maladie par poste.

On note donc, que ce soit au niveau des sommes des frais des soins médicaux ou de fréquences de sinistres , le poste pharmacie arrive en tête avec respectivement 33% et 40%.

La deuxième place est occupée par le poste ambulatoire avec 34% de fréquence de sinistres, et par le poste soins & Prothèses avec 19% du somme des frais de soins médicaux.

I.2 Analyse de la corrélation :

Afin de suspecter d'éventuelles redondances entre les variables quantitatives, il fallait procéder à une étude de corrélation. Cela a nécessité l'exportation de notre fichier initial de base de données vers le logiciel SPSS et la création d'autres variables à savoir :

- L'âge moyen au sien de chaque groupe **: Age_mean**
- Le montant complémentaire de remboursement de base **: Mont_cplmt_base**[1]
- Le code de l'âge moyen **: Agecode**
- Le code de la taille de chaque groupe **: Taillecode**
- Le code de plafond de remboursement de base **: Plafondcode**

Le ficher SPSS ainsi obtenu, nommé Base1, rassemble toutes les prestations effectuées par la compagnie en faveur de tous les bénéficiaires durant la période susmentionnée, il comportera environ 104000 observations. Un extrait de ce fichier est représenté par la figure 2.3 ci-dessous.

[1] Mont_cplmt_base : c'est la différence entre le montant réellement engagé et le montant de remboursement.

Base_1.sav [DataSet0] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1 : AASURV 2004 Visible: 20 of 20 Variables

	PRIXUNIT	FRAISREL	BASRMB	COEFACT	MNTRMB	MNTPLF	taille	age_mean	Mont_cplmt_base	Taillecode	agecode	Plafondcode
103203	0,00	500,00	500,00	100,00	500,00	50000,00	327	26,80	0,00	4	6	
103204	0,00	200,00	200,00	100,00	200,00	50000,00	327	26,80	0,00	4	6	
103205	0,00	3500,00	3500,00	100,00	3500,00	50000,00	327	26,80	0,00	4	6	
103206	0,00	238,90	238,90	100,00	238,90	50000,00	327	26,80	0,00	4	6	
103207	0,00	200,00	200,00	100,00	200,00	50000,00	327	26,80	0,00	4	6	
103208	0,00	600,00	600,00	100,00	600,00	50000,00	327	26,80	0,00	4	6	
103209	0,00	800,00	800,00	100,00	800,00	50000,00	327	26,80	0,00	4	6	
103210	0,00	53,90	53,90	100,00	53,90	50000,00	327	26,80	0,00	4	6	
103211	0,00	250,00	250,00	100,00	250,00	50000,00	327	26,80	0,00	4	6	
103212	0,00	900,00	900,00	100,00	900,00	50000,00	327	26,80	0,00	4	6	
103213	0,00	200,00	200,00	100,00	200,00	50000,00	327	26,80	0,00	4	6	
103214	0,00	400,00	400,00	100,00	400,00	50000,00	327	26,80	0,00	4	6	
103215	0,00	355,00	355,00	100,00	355,00	50000,00	327	26,80	0,00	4	6	
103216	0,00	200,00	200,00	100,00	200,00	50000,00	327	26,80	0,00	4	6	
103217	0,00	400,00	400,00	100,00	400,00	50000,00	327	26,80	0,00	4	6	
103218	0,00	180,85	180,85	100,00	180,85	50000,00	327	26,80	0,00	4	6	
103219	0,00	300,00	300,00	100,00	300,00	50000,00	327	26,80	0,00	4	6	
103220	0,00	465,90	465,90	100,00	465,90	50000,00	327	26,80	0,00	4	6	
103221	0,00	161,10	161,10	100,00	161,10	50000,00	327	26,80	0,00	4	6	
103222	0,00	200,00	200,00	100,00	200,00	50000,00	327	26,80	0,00	4	6	
103223	0,00	400,00	400,00	100,00	400,00	50000,00	327	26,80	0,00	4	6	
103224	0,00	200,00	200,00	100,00	200,00	50000,00	327	26,80	0,00	4	6	
103225	0,00	400,00	400,00	100,00	400,00	50000,00	327	26,80	0,00	4	6	
103226	0,00	200,00	200,00	100,00	200,00	50000,00	327	26,80	0,00	4	6	
103227	0,00	400,00	400,00	100,00	400,00	50000,00	327	26,80	0,00	4	6	
103228	0,00	655,00	655,00	100,00	655,00	50000,00	327	26,80	0,00	4	6	

Data View Variable View

SPSS Processor is ready

Figure 2.3 : Extrait de la base de données finale (Base1).

Les résultats issus de l'étude de corrélation entre les trois variables quantitatives les plus pertinentes ont donné, d'une manière générale, des coefficients de corrélation de Pearson très faibles et une corrélation non significative au seuil de 5%.

Le tableau 2.3 ci-dessous et le tableau 2.4 représenté en annexe 2, montrent aussi l'existence d'une forte corrélation entre les variables âge moyen et le plafond de remboursement, et des coefficients de corrélation faibles et négatifs entre les variables : âge moyen et la taille de groupe, et entre les variables plafond de remboursement et la taille de groupe.

Correlations

		MNTPLF	Taille	age_mean
MNTPLF	Pearson Correlation	1,000	-,106**	,728**
	Sig. (2-tailed)		,000	,000
	N	103899,000	103899	103899
Taille	Pearson Correlation	-,106**	1,000	-,061**
	Sig. (2-tailed)	,000		,000
	N	103899	103899,000	103899
age_mean	Pearson Correlation	,728**	-,061**	1,000
	Sig. (2-tailed)	,000	,000	
	N	103899	103899	103899,000

**. Correlation is significant at the 0.01 level (2-tailed).

Tableau 2.4 Analyse de corrélation entre les variables continues.

Aussi, l'analyse de corrélation de Kendall et de Spearman conforte les résultats obtenus (voir tableau 2.5 à l'annexe 2).

Ces résultants expliquent bien le fait que les plafonds de remboursement augmentent avec l'âge de l'assuré puisque celui-ci devient plus exposé au risque de maladies. Alors chaque fois que la taille du groupe est importante, la fréquence de sinistralité devient grande, et par suite les plafonds de remboursement seront réduits.

Les trois variables : Agecode, Taillecode et Plafondcode ont été obtenus suite de segmentations puis de codifications respectivement des trois variables : L'âge moyen au sein de chaque groupe, la taille de chaque groupe et le plafond de remboursement de base. Le chapitre 2 suivant, traitera toutes ces segmentations et les tests correspondants.

Chapitre II: Segmentation et codification des variables

Rappelons que le but de notre étude consiste à **modéliser le Ticket modérateur** de la consommation médicale, évalué par la différence entre les montants remboursés à l'assuré et les frais réellement engagés par celui-ci, en vue de déduire un tarif de base propre à chaque

groupe. Cela nécessite une sélection des variables les plus pertinentes qui expliquent la variable d'étude.

Ensuite, il faut procéder à une segmentation des variables tarifaires. Dans cette partie, l'objectif est de constituer des classes de risque homogènes de façon à ce que les assurés qui se retrouvent dans la même série de classes soient soumis au même tarif.

II.1. Importance de la segmentation :

Tout assureur est partagé entre mutualiser et sélectionner les risques. La segmentation est un moyen pour l'assureur de lutter contre l'anti-sélection ; un assureur qui ne segmente pas risque de perdre de bons risques au profit de celui qui segmente son portefeuille en offrant plus d'informations aux assurés. Les actuaires utilisent souvent des segmentations reposant sur les algorithmes des arbres de décision. Dans cette étude, nous avons utilisé l'algorithme de CHAID.

Cependant, toute bonne segmentation doit fournir des classes homogènes en intra (variabilité intra classe minimale) et hétérogènes en extra (variabilité inter classes maximale) du point de vue de la variable expliquée, et chaque classe doit contenir un nombre significatif d'observations.

II.2. Algorithme de CHAID :

Le but de l'algorithme de CHAID consiste à mesurer le degré de dépendance entre la variable d'intérêt qui est le critère de segmentation et les variables susceptibles de l'expliquer. Il s'agit de former des groupes en choisissant à chaque étape la variable la plus discriminante vis-à-vis de la variable expliquée de manière à générer à la fin un arbre de segmentation.

L'algorithme permet d'une part de sélectionner les variables les plus pertinentes qui expliquent le critère de segmentation, d'autre part de catégoriser les variables tarifaires en différentes modalités faciles à codifier.

Maintenant, nous allons présenter le test utilisé pour mesurer le degré de liaison entre la variable expliquée et les groupes formés. Nous le ferons pour deux groupes et généraliserons pour le cas de plusieurs groupes.

Soient P une population de taille n et $X^1, X^2, ...; X^K$, K variables explicatives définies sur P, M étant la moyenne générale de la population et σ^2 sa variance.

Notre population sera divisée en deux sous populations P_1 et P_2. M_1 et M_2 sont les sous moyennes et σ_1^2 et σ_2^2 les sous variances.

Soit X^j la jème variable explicative $X^j = [X_1^j, X_2^j, \dots, X_n^j]$.

On a : (1)

$$\boxed{\mathbf{M} = \frac{1}{n} * \sum_i \sum_j X_i^j = \frac{n_1}{n} * M_1 + \frac{n_2}{n} * M_2}$$

Avec : n_1 et n_2 sont les effectifs des deux groupes.

Et:

$$\boldsymbol{\sigma^2} = \left[\frac{n_1}{n} * \sigma_1^2 + \frac{n_2}{n} * \sigma_2^2\right] + [\frac{n_1}{n} * (M_1 - M) + \frac{n_2}{n} * (M_2 - M)]$$

La relation (1) peut aussi s'écrire sous la forme:

$$\boxed{\boldsymbol{\sigma^2} = \sigma^2(D) + \sigma^2(E)}$$

Avec:

σ^2 (D) : La variance dans les groupes.

σ^2 (E) : La variance entre les groupes.

Pour pouvoir séparer la population en deux groupes homogènes, le module de segmentation cherche à construire les groupes qui minimisent σ^2 (D) et maximisent σ^2(E).

Les hypothèses à vérifier sont les suivantes :

- H_0 : Les moyennes de Y sur chaque segment sont égales.
- H_1 : Les moyennes de Y sur chaque segment sont différentes.

Le test se base sur le calcul de la statistique :

$$\boxed{R = \frac{Variance\ interclasse}{Variance\ intraclasse}}$$

On a : $R \sim F(1, N-1)$; statistique de Fisher, où N est le nombre total d'observations.

On note que ce qui est valable pour deux groupes peut être généralisé à k groupes en remplaçant les variances par les matrices des variances-covariances.

II.3. Sélection des variables tarifaires :

L'application de l'algorithme de CHAID, via le logiciel SPSS, avec comme critère de segmentation la somme des tickets modérateurs c.à.d. les montants complémentaires de remboursement de base (**Mont_cplmt_base_sum**) accordés aux assurés d'un même groupe, a produit, comme illustre bien la figure 2.3, un arbre qui segmente le portefeuille de l'assureur en faisant apparaître un ensemble de variables tarifaires qui sont toutes significatives, à savoir :

- L'âge moyen dans chaque groupe : Agecode
- La taille de chaque groupe : Taillecode
- Le plafond de remboursement de base : Plafondcode

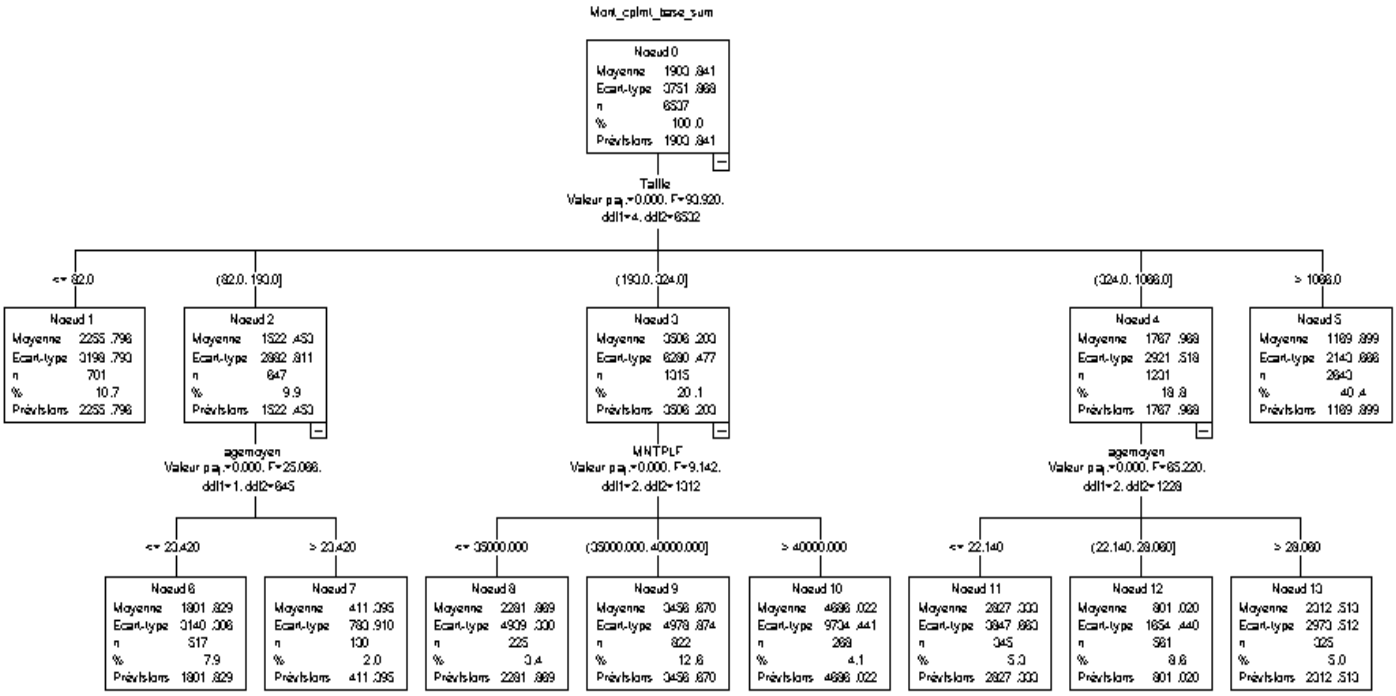

Figure 2.4 : Arbre de segmentation jusqu'au 3éme niveau pour le ticket modérateur.

II.4 Segmentation et Codification des variables tarifaires

Afin de catégoriser les variables tarifaires en vue de la modélisation GLM, nous avons appliqué l'algorithme de CHAID à chacune des trois variables tarifaires.

II.4.1 Segmentation et Codification de la variable Taille de groupe :

Nous présentons les modalités obtenues par l'algorithme du CHAID pour la variable taille de groupe renommée Taille ;

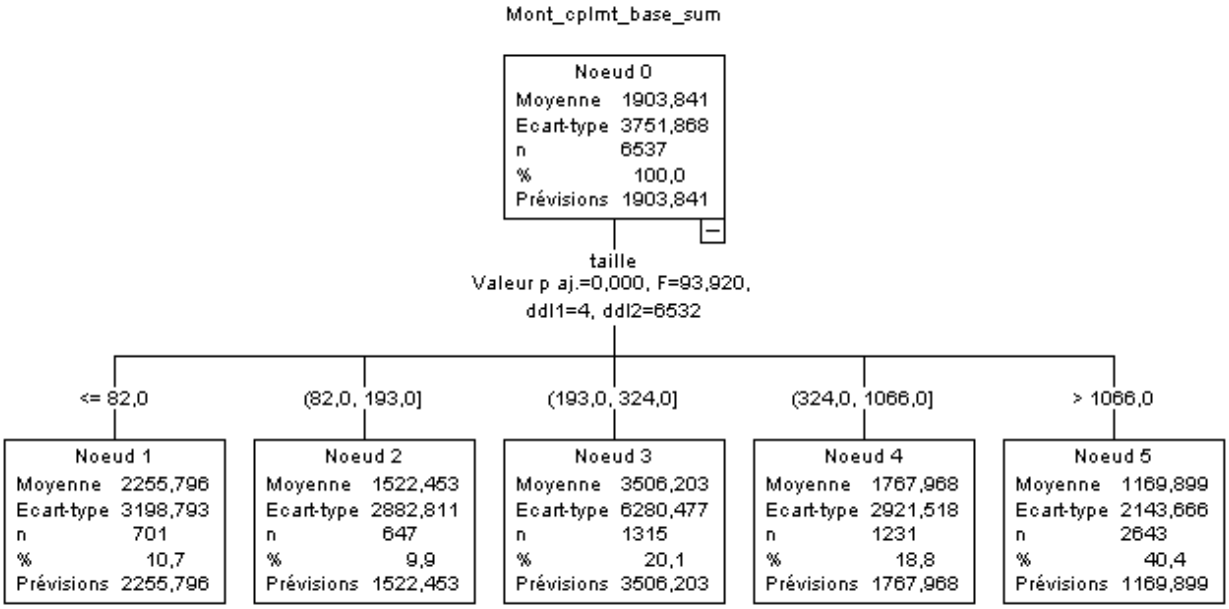

Figure 2.5 : Arbre de segmentation de la variable Taille de groupe.

Comme illustrent bien la figure 2.5et le Tableau2.6 des statistiques, cette segmentation a donné 5 modalités réparties comme suit :

Code	1	2	3	4	5
Modalité de taille du groupe	≤ 82	] 82 ; 193]	] 193 ; 324]	] 324 ; 1066]	≥ 1066
Moyenne (en DH) du Mont_cplmt_base_sum	2255,79	1522,45	3506,20	1767,96	1169,90
pourcentage	10,7%	9,9%	20,1%	18,8%	40,4%

Tableau 2.6: Les modalités et codes correspondants de la variable taille de groupe.

Récapitulatif des gains pour les noeuds

Noeud	N	Pourcentage	Moyenne
3	1315	20,1%	3506,2034
1	701	10,7%	2255,7957
4	1231	18,8%	1767,9684
2	647	9,9%	1522,4527
5	2643	40,4%	1169,8991

Méthode de développement : CHAID

Variable dépendante : Mont_cplmt_base_sum

Tableau 2.7 : Les statistiques issues de la segmentation de la variable Taille de groupe.

II.4.2 Segmentation et Codification de la variable âge moyen de groupe :

Les modalités obtenues par l'algorithme de CHAID pour la variable âge moyen au sein d'un même groupe renommée agemoyen sont représentées dans l'arbre de figure 2.6 ci-dessous.

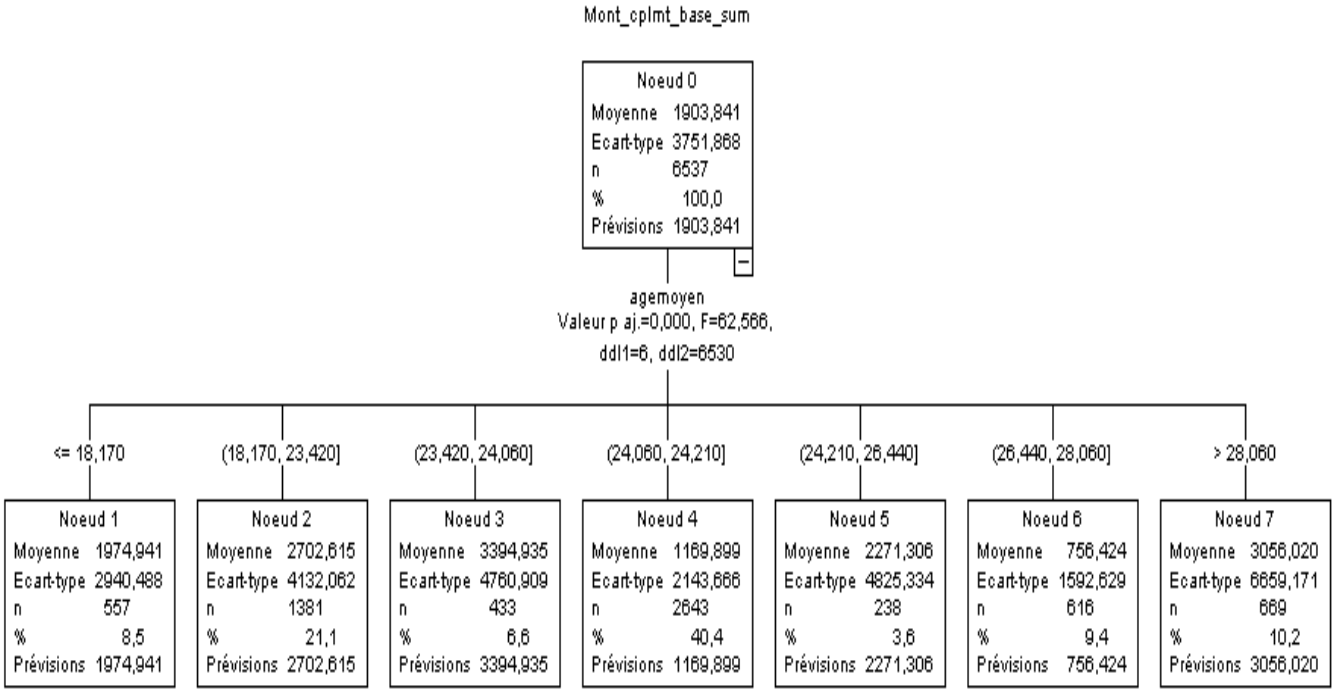

Figure 2.6 : Arbre de segmentation de la variable âge moyen de groupe.

Comme illustrent bien la figure 2.6 et le tableau2.9 des statistiques, cette segmentation a donné sept modalités réparties comme suit :

Code	1	2	3	4	5	6	7
Modalité d'âge moyen (ans)	≤ 18,17	]18,17 ; 23,42]	] 23,42 ; 24,06]	] 24,06 ; 24,21]	] 24,21 ; 26,44]	] 26,44; 28,06]	≥ 28,06
Moyenne (DH) du Mont_cplmt_base_sum	1974,94	2702,60	3394,93	1169,90	2271,30	756,42	3056
Pourcentage	8,5%	21,1%	6,6%	40,4%	3,6%	9,4%	10 ,2%

Tableau 2.8: Les modalités et codes correspondants de la variable âge moyen de groupe.

Récapitulatif des gains pour les noeuds

Noeud	N	Pourcentage	Moyenne
3	433	6,6%	3394,9348
7	669	10,2%	3056,0205
2	1381	21,1%	2702,6154
5	238	3,6%	2271,3062
1	557	8,5%	1974,9405
4	2643	40,4%	1169,8991
6	616	9,4%	756,4241

Méthode de développement : CHAID

Variable dépendante : Mont_cplmt_base_sum

Tableau 2.9: Les statistiques issues de la segmentation de la variable âge moyen.

II.4.3 Segmentation et Codification de la variable Montant Plafond de remboursement :

La segmentation de la variable Montant Plafond de remboursement noté : MNTPLF, obtenue par l'algorithme de CHAID, est représentée dans l'arbre de figure 2.7 ci-dessous.

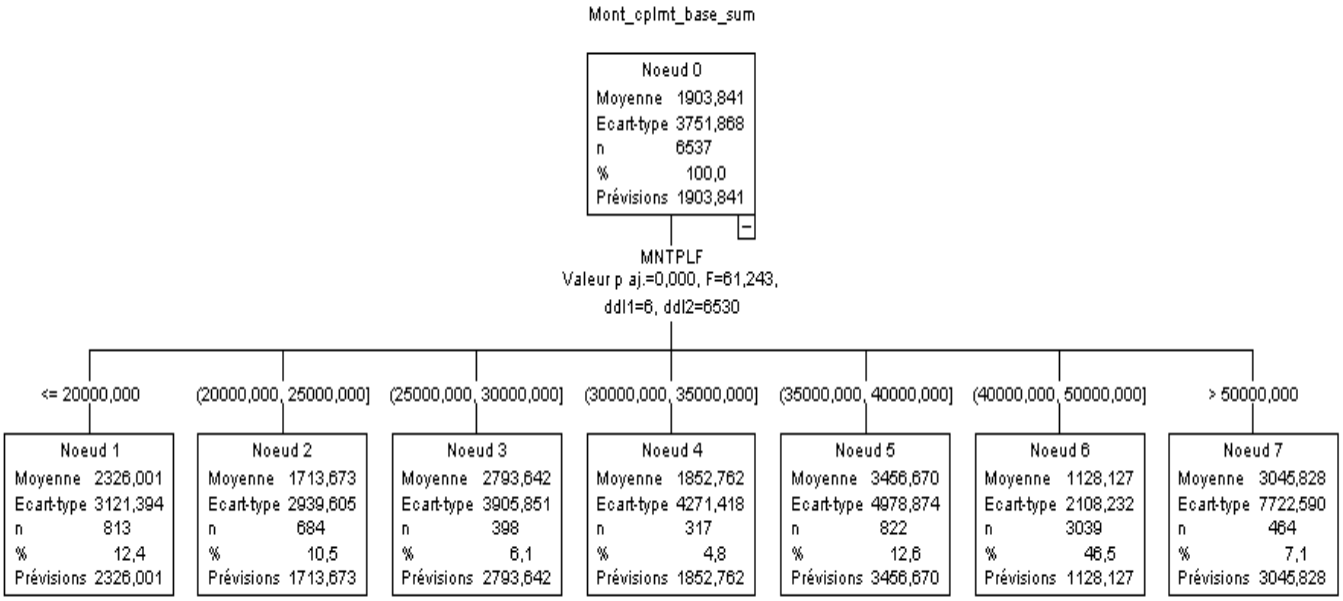

Figure 2.7: Arbre de segmentation de la variable montant plafond de remboursement.

La figure 2.7 et le tableau 2.11 des statistiques, issues de cette segmentation a donné sept modalités réparties comme suit :

Code	1	2	3	4	5	6	7
Modalité de montant plafond de remboursement	≤ 20000	] 20000 ; 25000]	] 25000 ; 30000]	] 30000 ; 35000]	] 35000 ; 40000]	] 40000; 50000]	≥ 50000
Moyenne du Mont_cplmt_base_sum (DH)	2326	1713,67	2793,64	1852,76	3456,67	1128,87	3045,82
Pourcentage	12,4%	10,5%	6,1%	4,8%	12,6%	46,5%	7,1%

Tableau 2.10 : Les modalités et codes correspondants de la variable montant plafond de remboursement.

Récapitulatif des gains pour les noeuds

Noeud	N	Pourcentage	Moyenne
5	822	12,6%	3456,6705
7	464	7,1%	3045,8281
3	398	6,1%	2793,6417
1	813	12,4%	2326,0011
4	317	4,8%	1852,7617
2	684	10,5%	1713,6726
6	3039	46,5%	1128,1265

Méthode de développement : CHAID

Variable dépendante : Mont_cplmt_base_sum

Tableau 2.11 : Les statistiques issues de la segmentation de la variable montant plafond de remboursement.

Ainsi, les différentes modalités issues de segmentations de chacune des trois variables ne puiseraient être utilisées qu'après avoir testé leurs différentiation. D'où le test de Kruskal Wallis.

II.5. Test de Kruskal Wallis:

Afin de s'assurer que les segments obtenus pour chaque variable sont significativement différents, on procède au test de kruskal Wallis qui consiste à vérifier l'égalité des montants complémentaires de remboursement de base **(Mont_cplmt_base_sum)** moyennes Y (le critère de segmentation) des classes ainsi formées. Il s'agit de tester les hypothèses suivantes :

- H_0 : Les moyennes de Y sur chaque classe sont égales.
- H_1 : Les moyennes de Y sur chaque classe sont différentes.

Pour comprendre la construction de la statistique de kruskal wallis qui converge vers une chi- Deux de degrés de liberté (1-α) où α est le niveau de signification du test, le lecteur pourra se référer à l'ouvrage de MUSTAPHA BENYAKLEF, *« Probabilités et statistique mathématique Tome II »*.

Le tableau2.9 ci-dessous récapitule les p-values obtenues pour les différentes variables.

Test[a,b]

	Taillecode	PLafondcode	agecode
Khi-deux	2877,694	3034,431	3173,895
ddl	2859	2859	2859
Signification asymptotique	,009	,011	,000

a. Test de Kruskal Wallis

b. Critère de regroupement : Mont_cplmt_base_sum

Tableau 2.12: Résultats du test de kruskal Wallis pour les 3 segmentations effectuées.

On note que toutes les p-values sont inferieures à 5%, donc on rejette l'hypothèse H_o pour chaque variable tarifaire. Et on conclut que les classes formées par l'algorithme de CHAID sont bien différentes.

Remarques :

i/ On aurait dû segmenter ces trois variables en prenant comme critère de segmentation la fréquence des sinistres comme étant elle aussi une variable de sinistralité, mais ces segmentations ont donné des classes qui ne sont pas significativement distinctes. Dans l'annexe 3 (figure 2.8), on représentera l'arbre de segmentation du portefeuille jusqu'au 3^{éme}niveau pour le ticket modérateur en prenant la fréquence des sinistres comme critère de segmentation.

ii/ Il est à noter aussi, que dans le but de mener une analyse descriptives plus pertinente des trois variables tarifaires, notamment en ce qui concerne la taille et l'âge moyen de chaque groupe, il fallait porter de légères modifications aux modalités de ces deux variables, ce qui nous a permis de mieux les décrire et de mieux les analyser, et ce qui fera l'objet du chapitre suivant.

Chapitre III : Description des variables tarifaires

III.1 La variable taille du groupe :

La répartition des dépenses sanitaires et de fréquence de sinistralité selon les modalités de la variable Taille est représentée dans le tableau 2.13 et dans les figures 2.9 et 2.10 ci-dessous :

Poste	Somme des Frais réels en (DH)	Somme des Frais réels en %	Fréquences	Fréquences en %
1 : Ambulatoire	6277676,44	13%	35095	34%
2 : Pharmacie	15476391,74	33%	41287	40%
3 : Hospitalisation	5381521,23	11%	3398	3%
4 : Analyse & Radios	5727896	12%	12555	12%
5 : Soins & Prothèses	8935592,23	19%	6598	6%
6 : Optique	4360827,69	9%	4527	4%
7 : Autres	1260778,79	3%	434	0%
Total	47420684,12	100%	103894	100%

Tableau 2.3 : Statistiques descriptives des différents Postes de maladie

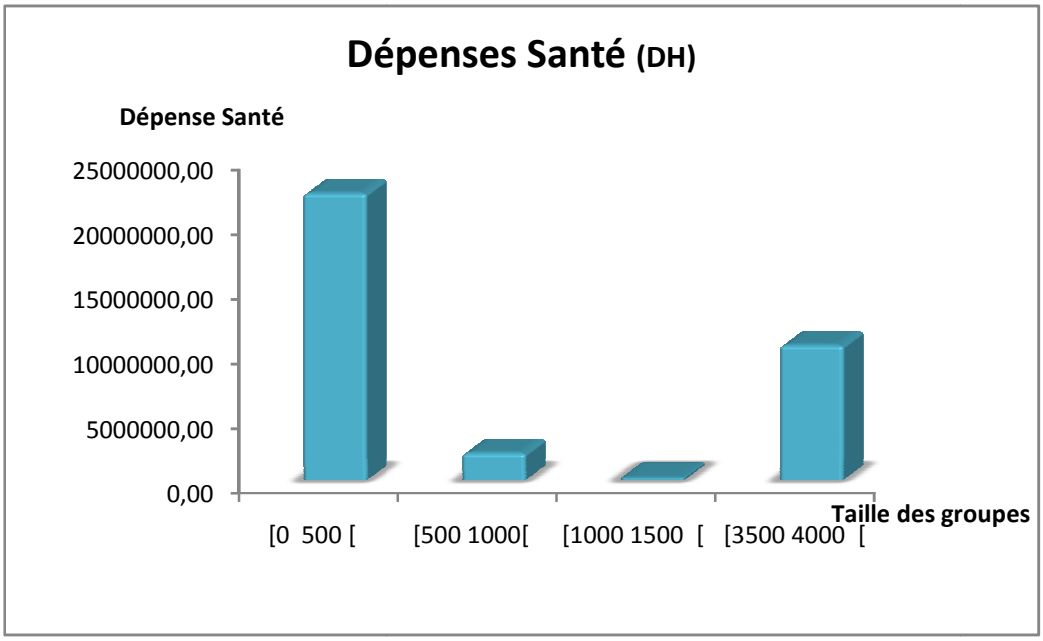

Figure 2.9 : Variation des dépenses sanitaires.

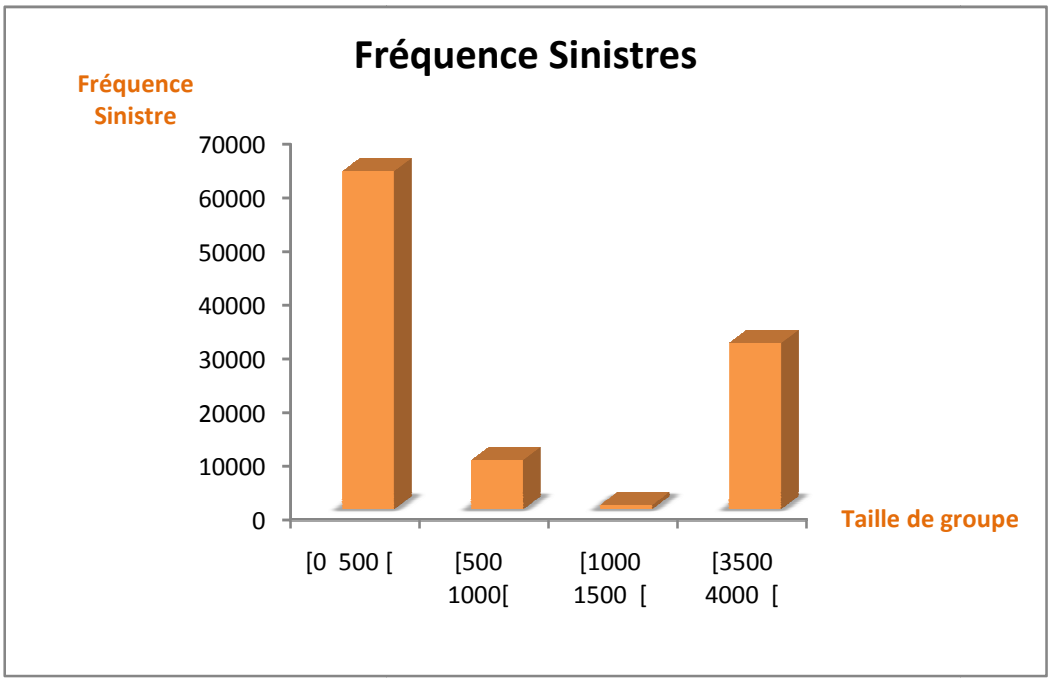

Figure 2.10: Variation de fréquence de sinistres.

III.2 La variable âge moyen du groupe:

La répartition des dépenses sanitaires et de fréquence de sinistralité selon les modalités de la variable âge moyen du groupe est représentée dans le tableau2.14 et dans les figures 2.11 et 2.12 ci-dessous :

Age Moyen(ans)	Dépenses Santé (DH)	Fréquence Sinistres
[0 20[	3266373,20	11112,74
[20 40[	31 711045,04	92781,11
[40 90[	1284,00	2,01

Tableau 2.14 : Répartition des dépenses sanitaires et de fréquence de sinistralité de la variable âge moyen du groupe.

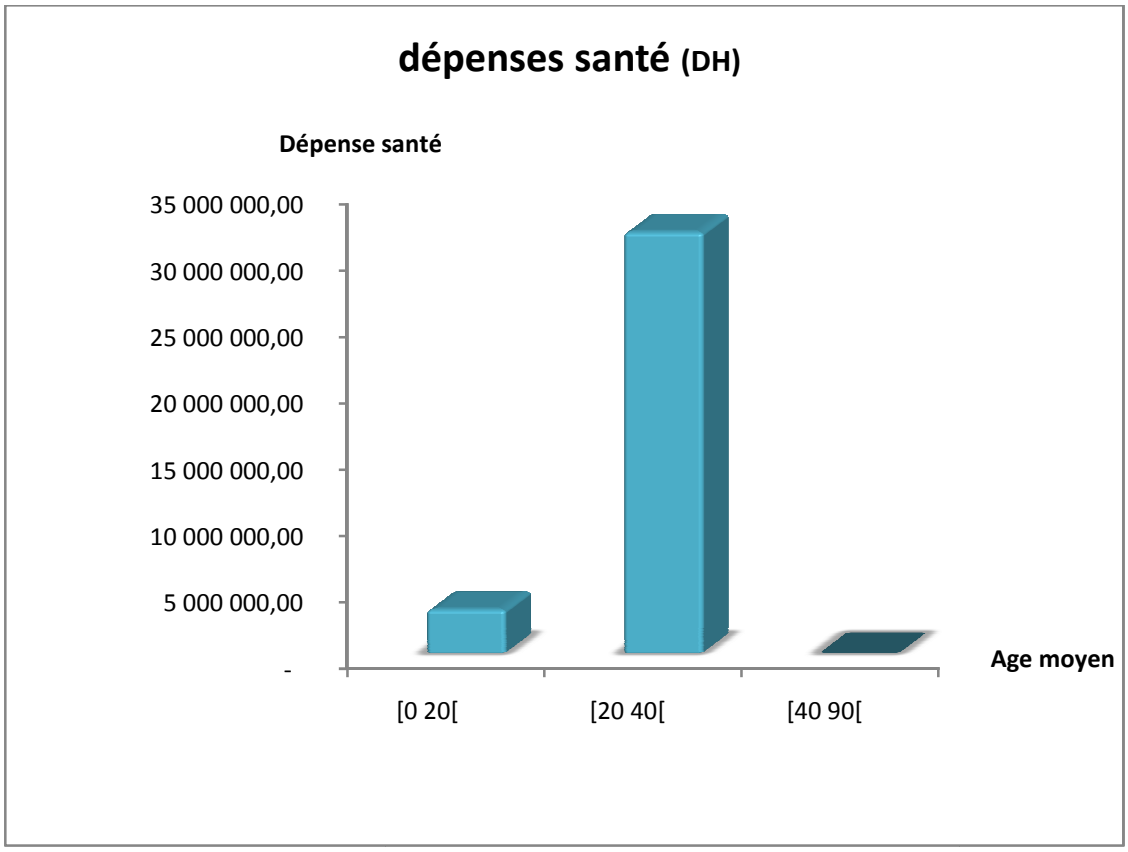

Figure 2.11: Variation des dépenses sanitaires.

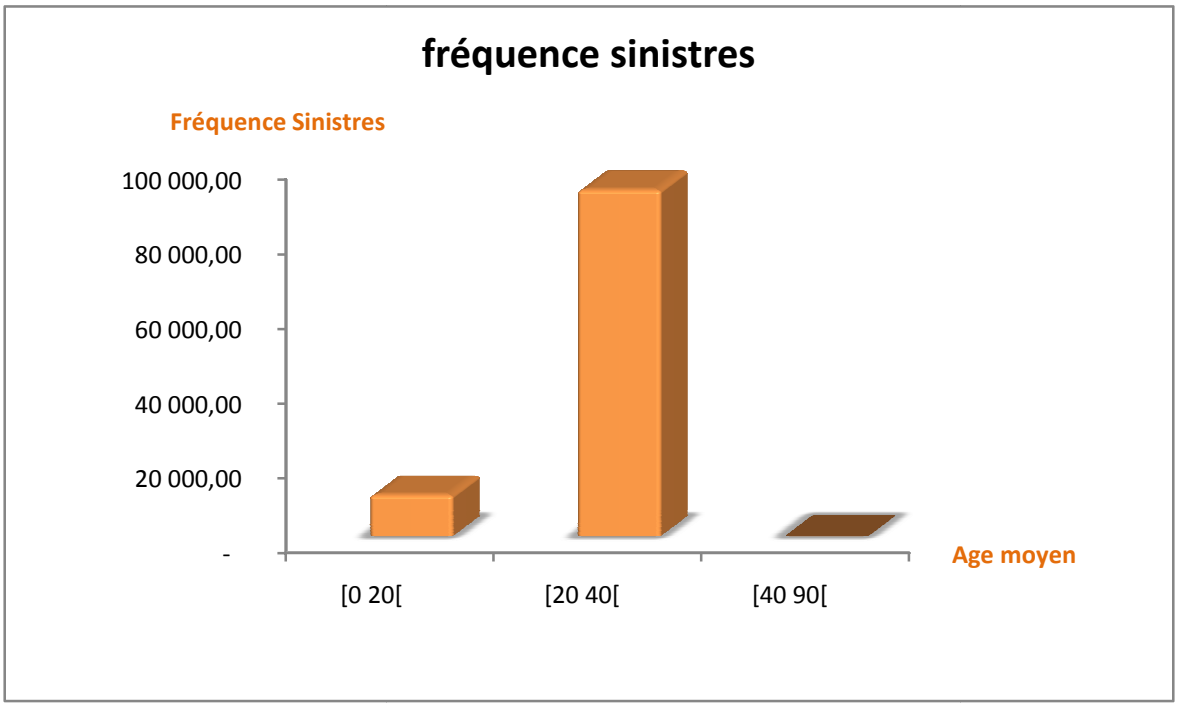

Figure 2.12: Variation de fréquence de sinistres.

III.3. La variable montant plafond de remboursement :

La répartition des dépenses sanitaires et de fréquence de sinistralité selon les modalités de la variable montant plafond de remboursement est représentée dans le tableau 2.15 et dans les figures 2.13 et 2.14 ci-dessous :

Plafond Remboursé	Dépenses Santé (DH)	Fréquence Sinistres
<=20000	2326	813
] 20000 25000]	1713,67	684
] 25000 30000]	2793,64	398
] 30000 35000]	1852,76	317
] 35000 40000]	3456,67	822
] 40000 50000]	1128,12	3039
>= 50000	3045,82	464

Tableau 2.15: Répartition des dépenses sanitaires et de fréquence de sinistralité de la variable Plafond de remboursement.

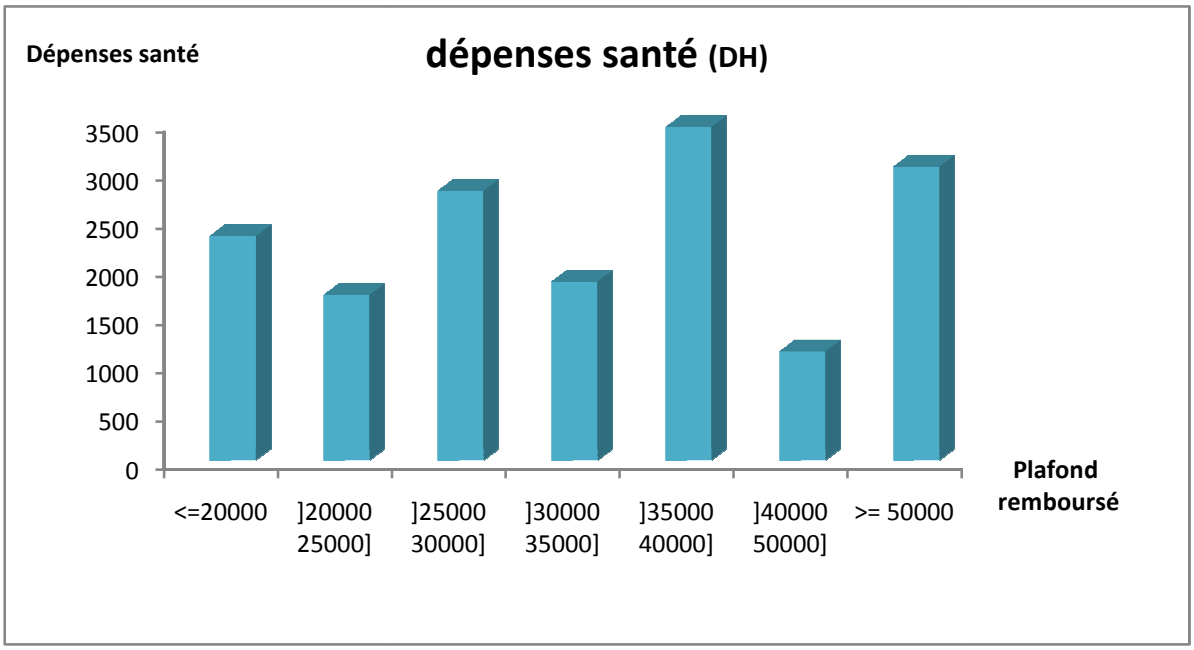

Figure 2.13 : Variation des dépenses sanitaires.

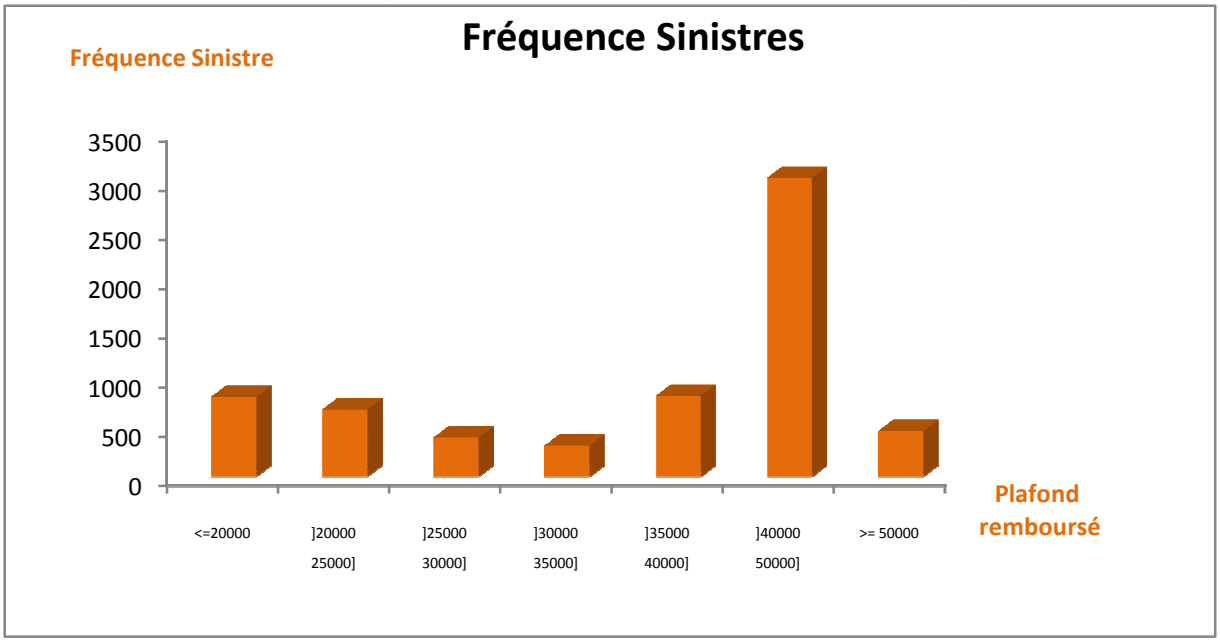

Figure 2.14: Variation de fréquence de sinistre

Partie III

Modèles de Tarification à Priori

Le but de cette partie est de procéder à une tarification à priori de chaque contrat groupe. Il s'agit de déterminer la prime pure de chaque police, c'est-à-dire le montant à hauteur duquel elle doit cotiser pour être indemnisé des sinistres causés au niveau de chaque poste maladie. Et la cotisation du groupe est la somme des primes dégagées dans tous les postes auxquels il a souscrit une assurance complémentaire. Pour chaque poste, nous envisageons deux méthodes de tarification.

La première approche est basée sur la modélisation du nombre de sinistres des polices en l'ajustant par une loi de probabilité adéquate et une régression GLM sur les caractéristiques propres aux polices (taille, âge moyen, plafond). Ensuite, on estime le coût moyen des sinistres par un ajustement graphique Q-Q plot. Ainsi la prime pure de chaque contrat groupe dans le poste est égale à son nombre moyen de sinistres multiplié par le cout moyen d'un sinistre.

La deuxième approche repose sur la modélisation logistique de la fréquence des sinistres et une estimation de la charge annuelle moyenne d'une police par la méthode graphique du Q-Q plot. Ainsi, la prime pure de chaque contrat groupe dans le poste est égale à sa probabilité de causer au moins un sinistre sur un an multipliée par sa charge annuelle moyenne de sinistralité.

Dans un premier temps, nous procédons à la modélisation des fréquences, et ensuite celle des couts moyens.

Chapitre I : Notion d'assurance de groupe.

Notre étude consiste à la modélisation de la prime pure en vue d'une assurance maladie complémentaire d'un contrat groupe ayant plusieurs bénéficiaires. On ne fait pas de distinction entre les sinistres provenant des différents adhérents, mais on les mets à l'actif de la police du groupe auquel ils appartiennent. Et dans le cadre de l'AMO, ce dernier peut être une administration publique, une entreprise privée ou un cabinet à profession libérale.

Si nous désignons par Ni le nombre de sinistres causés par l'adhérent i d'une police j donnée et admettons que les montants des sinistres de chaque adhérent sur toute la période de référence soient des variables aléatoires X_1, X_2…….indépendantes et identiquement distribuées de fonction de répartition F, nous pouvons formaliser mathématiquement un modèle collectif d'assurance sur la base de quelques hypothèses.

Hypothèses du modèle :

- Les variables aléatoires Xi sont i.i.d . Elles sont indépendantes du nombre des sinistres Ni.
- La fonction de répartition commune des Xi est telle que F(0)=0.
- Les variables Ni, X_1, X_2….ont leurs moments d'ordre 1 et 2 finis, c'est-à-dire que : E(Xi) , E(Ni), V(Xi) et V(Ni) sont finis.

Si l'on désigne par N et S les variables aléatoires correspondant respectivement au nombre de sinistres et au montant des sinistres touchant la police j, on pourrait écrire :

$N=\sum_i N_i$ et $S=\sum_i X_i$ avec S = 0 si N = 0.

L'espérance mathématique de S s'écrit E(S) = E(N).E(X). Elle s'interprète comme le produit du nombre moyen de sinistres de la police par le coût moyen d'un sinistre. Elle équivaut à la prime pure correspondante à la police.

La modélisation des variables aléatoires N et S en vue d'un tarif collectif requiert l'identification de caractéristiques de sinistralité propres au contrat groupe.

Partant des informations présentes dans notre base de données et relatives à chaque police, nous avons fait appel à la taille du groupe, son âge moyen et son plafond annuel comme variables censées expliquer sa sinistralité.

Chapitre II: Présentation de la première approche de tarification à priori.

Introduction :

L'objectif de ce chapitre est la détermination d'une prime pure basée sur la première approche dans chaque poste maladie. Nous présentons explicitement la démarche avec le poste pharmacie, et les détails des résultats obtenus pour les autres postes sont disponibles dans les annexes.

II .1 Modélisation de la fréquence :

Parmi les lois discrètes usuelles dans la modélisation actuarielle de la fréquence des sinistres, les lois de Poisson et binomiale négative sont celles qui fournissent très souvent de meilleurs ajustements. Dans notre contexte, les indicateurs statistiques d'ajustement tels que la déviance ou le critère d'Akaike (AIC) obtenus dans les modèles de poisson dépassent largement ceux qui sont obtenus avec les modèles de la binomiale négative. Le meilleur ajustement étant celui qui minimise ces indicateurs, la binomiale négative avec une régression GLM sur les caractéristiques des polices assurées a été retenue. Et une illustration sera exposée dans le poste pharmacie.

II.1.1 Bref aperçu sur les modèles linéaires généralisés :

Nous allons commencer par un bref aperçu théorique sur les modèles linéaires généralisés. En fait, les actuaires manipulent souvent des variables explicatives catégorielles (sexe, région, âge..), le nombre de sinistres qui est une variable discrète, le coût des sinistres ne pouvant pas prendre de valeurs négatives. Cela traduit donc les limites du modèle de régression multiple $y = X\beta + \varepsilon$ qui s'appuie sur le fait que la variable d'intérêt y soit normalement distribuée.

Les modèles GLM constituent une extension aux modèles de régression multiple. Ils permettent de modéliser toute variable réponse, en plus de la loi normale appartenant à la

famille de lois exponentielles (normal, gamma, exponentiel…), c'est-à-dire dont la fonction densité est de la forme :

$$f\ (y\ /\theta,\phi) = \exp\left\{\frac{y\,\theta - b(\theta)}{\phi} + c(y\ ,\phi)\right\}\quad x\in S\ \textbf{(*)}.$$

Où le support S est un sous ensemble de ℕ ou de ℝ ; θ est appelé paramètre naturel et ϕ le paramètre de dispersion.

Le tableau ci-dessous donne quelques modèles de régression usuels avec leurs fonctions lien :

Distribution	**Normale** **N** (μ,σ^2)	**Poisson** **P** (λ)	**Bernoulli** **B** (1, μ)	**Gamma** **G** (μ, η)
ϕ	σ^2	1	1	$\frac{1}{\eta}$
$\theta(\mu)$	M	Ln (μ)	Logit (μ)	$-\frac{1}{\mu}$
$b(\theta)$	$\frac{\theta^2}{2}$	$\exp(\theta)$	Ln $(1+\exp(\theta))$	- ln (-θ)
Lien canonique	**Identité**	**ln**	**Logit**	**Inverse**

Tableau 3.1 : Tableau récapitulatif de quelques modèles GLM usuels.

Considérons une suite de variables aléatoires $Y_1 \ldots\ldots Y_n$ indépendantes et identiquement distribuées i. i. d, admettant toutes la densité (*).

Ainsi, la fonction de vraisemblance s'écrit :

$$f\ (y\ /\theta,\phi) = L(\theta,\phi)\ = \prod_{i=1}^{n} f\ (y_i\ /\theta_i\ ,\phi)\ =\ \text{Exp}\ \left(\frac{\sum_{i=1}^{n} y_i\theta_i - \sum_{i=1}^{n} b(\theta_i)}{\phi} + \sum_{i=1}^{n} c(y_i,\phi)\ \right).$$

Avec les θ_i sont en fonction de (p+1) paramètres β_0, …, β_p à estimer par la maximisation de la log vraisemblance définie par :

$$\mathcal{L}\ (\theta(\beta)/y,\phi) = \sum_{i=1}^{n} ln^{f\ (y_i\ /\theta_i\ ,\phi)} = \frac{\sum_{i=1}^{n} y_i\theta_i - \sum_{i=1}^{n} b(\theta_i)}{\phi} + \sum_{i=1}^{n} c(y_i,\phi).$$

Il s'agit de rechercher les estimateurs du maximum de vraisemblance $\beta_0, \dots, \beta_p$ qui vérifient les conditions de premier ordre c'est-à-dire $U_j = 0$ avec :

$$U_j = \frac{\partial}{\partial \beta_j} \mathcal{L}(\theta(\beta)/y, \phi) = \sum_{i=1}^{n} \frac{\partial}{\partial \beta_j} ln f\,(y_i / \theta_i, \phi) = \sum_{i=1}^{n} \frac{\partial}{\partial \beta_j} \left(\frac{y_i \theta_i - b(\theta_i)}{\phi} + c(y_i, \phi) \right).$$

La résolution numérique de ces équations de vraisemblance $U_j=0$ se fait par la méthode de newton Raphson.

Si l'on pose $E[Y_i] = \mu_i$, un modèle GLM se résume à trois éléments essentiels :

- Les variables à expliquer $Y_1, \dots, Y_n$ dont les densités sont de la famille de loi exponentielle.
- Le prédicteur linéaire $\eta_i = x_i^t \beta$ avec des paramètres $\beta = (\beta_0, \dots, \beta_p)^t$ appartenant à un ouvert non vide de $\mathbb{R}^{p+1}$ et de variables explicatives $X = (x_1, \dots, x_n)^t$: la matrice X de dimension n * (p+1) est supposée de plein rang, c'est-à-dire que la matrice $X^t X$ est inversible.
- La fonction de lien canonique g telle que $g(\mu_i) = x_i^t \beta = \eta_i$ qui lie le prédicteur linéaire $\eta_i = x_i^t \beta$ à la moyenne μ_i de Y_i.

Voici une brève présentation des lois utilisées dans la modélisation GLM de la fréquence de sinistres :

- **Loi de poisson** :

Si une variable aléatoire N~Poisson(λ), sa fonction densité s'écrit :

$$\mathbf{P}\,(N=k) = \text{Exp}\,(-\lambda)\frac{\lambda^k}{k!}, \quad k \in N.$$

Les études actuarielles montrent empiriquement que la distribution de poisson donne de très bons ajustements du nombre de sinistres dans les portefeuilles homogènes.

- **Loi binomiale négative** :

Si une variable aléatoire N~binomiale négative (α, q) , sa fonction densité s'écrit :

$$\mathbf{P}\ (N{=}k) = \binom{\alpha + k - 1}{k} q^{\alpha}(1-q)^{k}, \quad k \in N.$$

Elle ajuste mieux les portefeuilles hétérogènes ; nous l'avons retenue dans tous les modèles GLM utilisés dans ce travail.

II-1-2 Application du modèle GLM à la modélisation de la fréquence :

Considérons une police assurée i disposant des caractéristiques (taille, âge moyen, plafond) résumées dans le vecteur $X_i = (x_1, \dots, x_n)^t$. Soit N_i le nombre de sinistres causés annuellement par la police.

Nous cherchons à estimer son nombre moyen de sinistres $\mu_i = E\ (N_i/X_i)$. Ainsi, en choisissant judicieusement comme fonction lien le logarithme népérien, nous obtenons le modèle multiplicatif suivant :

$$\ln(\mu_i) = \ln[E(N_i/X_i)] = x_i^t\beta \underset{\text{ou bien}}{\Longleftrightarrow} \mu_i = E\ (N_i/X_i) = \mathrm{Exp}(x_i^t\beta).$$

Il ne reste plus qu'à estimer les paramètres β du modèle.

II-1-3 Présentation du modèle pour le poste pharmacie :

Nous allons commencer par une analyse de corrélation entre les différentes variables explicatives. Cela se traduit par le tableau ci-dessous :

Corrélations

		agecode	Taillecode	PLafondcode
agecode	Pearson Correlation	1,000	**,543**	**,441**
	Sig. (2-tailed)		,000	,000
	N	129,000	129	129
Taillecode	Pearson Correlation	,543	1,000	,586
	Sig. (2-tailed)	,000		,000
	N	129	129,000	129
PLafondcode	Pearson Correlation	,441	**,586**	1,000
	Sig. (2-tailed)	,000	,000	
	N	129	129	129,000

Tableau 3.2: analyse de corrélation dans le poste pharmacie.

On note une légère corrélation entre la variable taillecode (un coefficient de corrélation de Pearson supérieur à 0.5) avec chacune des variables plafondcode et agecode.

La taille des groupes assurés est une variable sinistralité importante et facilement accessible comme information pour l'assureur. C'est pourquoi nous avons convenu à ce qu'elle figure dans tous les modèles suspectés. Ainsi, nous avons essayé toutes les combinaisons de la variable taille avec les deux autres aussi bien pour les ajustements de la binomiale négative que ceux de la loi de poisson.

Ensuite, en utilisant principalement la déviance des modèles comme critère de meilleur ajustement, et puis celui d'Akaike (AIC), nous avons choisi le modèle adéquat.

II-1-3-1 Choix du modèle sur la base de la qualité de l'ajustement.

L'un des outils pour analyser la qualité d'un modèle de régression est de s'appuyer sur la Déviance. On définit la qualité d'un modèle en prenant comme référence le modèle saturé ayant autant de paramètres que d'observations, ce qui lui confère une meilleure description des données.

Mais ce modèle étant redondant, en pratique on prend un modèle pour lequel le nombre de paramètres p est inferieur au nombre d'observations n.

On compare le modèle estimé avec le modèle saturé sur la base de la déviance D des logs vraisemblances l et l_{sat} avec $D = -2\ (l - l_{sat})$ comme étant le logarithme népérien du carré du rapport des vraisemblances.

En principe, une valeur de D très grande indique une mauvaise qualité du modèle. Mais en pratique, pour juger de la qualité d'un modèle on compare la déviance avec le nombre de degrés de liberté. Plus le rapport est plus petit ou proche de l'unité, plus le modèle peut être jugé comme satisfaisant.

Le tableau ci-dessous présente des modèles avec leurs déviances :

modèle	ajustement	critère	valeur	Degrés de liberté (ddl)	Valeur/ddl	AIC
Plafondcode & taillecode.	Binomiale négative	déviance	107.147	118	0.908	1585.49
Agecode & taillecode	Binomiale négative	déviance	114.551	119	0.963	1587.034
taillecode	Binomiale négative	déviance	130.974	124	1.056	1574.159
Agecode , Taillecode & plafondcode	poisson	déviance	5280.15	113	46.727	6186.262
taillecode	poisson	déviance	10557.896	124	85.144	11399.55

Tableau 3.3: Tableau récapitulatif des modèles pour la fréquence du poste pharmacie.

Ces résultats confortent le fait que la loi de poisson ne soit pas adaptée à nos données. On note des rapports déviance/degrés de libertés largement supérieurs à 1 et des critères d'akaike (AIC) de très grande valeur. Cela vaut aussi pour tous les autres postes maladie (voir annexes).

Ainsi, le meilleur modèle pour ce poste pharmacie est un ajustement de la binomiale négative qui retient la taille comme seule variable significative. On remarque pour ce modèle un rapport déviance/dl égal à 1.056 plus proche de 1 et un critère AIC minimal.

Il s'est d'ailleurs avéré que pour cette première approche de modélisation de la fréquence, la taille est la seule variable significative pour tous les autres postes (voir les annexes 4 à 10).

Nous allons ensuite procéder à une analyse de type III pour tester le niveau de significativité de la variable taille dans le modèle retenu.

II-1-3-2 Test de significativité de Wald :

Pour tester la significativité du coefficient β_k d'un facteur k dans un modèle, le test de Wald est souvent utilisé.

Soit le test suivant :

$$H_0 : \beta_k = 0 \quad \text{contre} \quad H_1 : \beta_k \neq 0$$

Ainsi la statistique de Wald s'écrit: $W = \dfrac{\hat{\beta}_k}{\hat{s}\left(\hat{\beta}_k\right)}$, avec $\hat{s}\left(\hat{\beta}_k\right)$: l'écart type estimé de $\hat{\beta}_k$. Sous H0 : W suit approximativement la loi normale N (0 ,1).

<u>Règle de décision :</u>

Au niveau de signification α : On rejette $H_0 \Leftrightarrow W > \phi(1-\alpha)$, avec ϕ: la fonction réciproque de la loi normale ; ou bien on rejette H_0 lorsque la P-value du test est inférieure à α.

Dans le cas de notre modèle, le test de Wald donne un niveau de signification de 0.00 pour la variable taillecode comme en illustre le tableau ci-dessous :

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	2328,652	1	,000
Taillecode	181,815	4	,000

<u>Tableau 3.4</u>: Analyse de type III de la variable taillecode.

Nous déduisons de ce tableau que la taille contribue significativement à l'explication de la fréquence des sinistres au seuil de 5%.

II-1-3-3 Estimation des paramètres du modèle :

L'estimation des paramètres du modèle a donné les résultats ci-dessous :

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	7,763	,4473	6,886	8,639	301,168	1	,000	2351,200	978,452	5649,883
[Taillecode=1]	-3,770	,4637	-4,679	-2,861	66,074	1	,000	,023	,009	,057
[Taillecode=2]	-2,761	,5004	-3,742	-1,780	30,436	1	,000	,063	,024	,169
[Taillecode=3]	-1,336	,5002	-2,316	-,356	7,133	1	,008	,263	,099	,701
[Taillecode=4]	-1,276	,5125	-2,280	-,271	6,194	1	,013	,279	,102	,763
[Taillecode=5]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									
(Negative binomial)	1									

Dependent Variable: Fréquence
Model: (Intercept), Taillecode

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Tableau 3.5 : Estimation des paramètres pour la fréquence du poste pharmacie.

Le test de Wald indique que toutes les modalités de la variable taillecode sont significatives au seuil de 5%. Dans un modèle GLM, chaque modalité est introduite sous forme binaire. Elle vaut 1 si la police y appartient et 0 sinon.

Les coefficients multiplicatifs Exp (β) s'interprètent par rapport à une classe de référence constituée des modalités dont le coefficient multiplicatif vaut 1. IL s'agit ici de la modalité taille code 5 regroupant les polices dont la taille est supérieure à 1066 adhérents.

Le risque se définit comme une fréquence médicale moyenne supérieure à celle de la classe de référence. Donc si une modalité est associée à un coefficient supérieur à 1, cela signifie une majoration par rapport à la fréquence médicale pharmaceutique de la classe de référence, et donc une aggravation de la sinistralité.

De même si une modalité est associée à un coefficient inférieur à 1, cela signifie une minoration par rapport à la fréquence médicale de la classe de référence, et donc une réduction de la sinistralité.

Avec les paramètres ainsi estimés, nous pouvons écrire l'expression du nombre moyen de sinistres pour une police i donnée :

$$\boldsymbol{\mu_i = E(N_i/X_i) = e^{(\hat{\beta}_0)} . \prod_{j=1}^{5} e^{taillecodej(i).\hat{\beta}_j}}$$

Avec $taillecode_j(i)$: la valeur de la $j^{éme}$ modalité de la variable taillecode pour la police i.

II.2 Modélisation du coût moyen d'un sinistre :

La modélisation du coût moyen des sinistres passe par l'ajustement des montants de sinistre par une distribution de probabilité adéquate avec ses paramètres estimés. Toutefois, le modèle retenu par l'actuaire reste lié à la qualité des bases de données historiques. Il doit pouvoir l'améliorer s'il ne s'adapte plus à la réalité.

Il apparaît souvent dans les modèles en actuariat que la loi Gamma ajuste mieux les petits montants (< 5000dhs par exemple), la loi Lognormal convient pour les montants moyens et Pareto pour les grands montants (>100000dhs par exemple). Cela pourrait aussi se justifier par l'épaisseur des fonctions queue [2] correspondant à ces lois.

La méthode graphique du Q-Q plot est très instructive pour le choix du modèle. Pour la plupart des modèles paramétriques, les quantiles observés $F_{\theta}^{-1}(p)$ sont en relation linéaire avec les quantiles correspondant à la loi $F_{\theta*}^{-1}$.

En d'autres termes, le graphique Q-Q plot met en relation linéaire à travers un nuage de points les quantiles observés et les quantiles expectés par une loi d'ajustement.

Dans ce travail, nous avons utilisé la procédure Q-Q plot du logiciel SPSS qui fournit le graphique Q-Q plot avec les paramètres estimés de la loi.

II.2.1 Présentation du modèle pour le poste pharmacie :

II.2.1.1 Q-Q plot des lois suspectées

Nous présentons ici les Q-Q plots des lois suspectées pour les montants complémentaires de sinistres dans le poste pharmacie.

2 la fonction queue d'une loi est égale à 1-F(x) où F(x) est sa fonction de répartition. Elle mesure la dangerosité des sinistres.

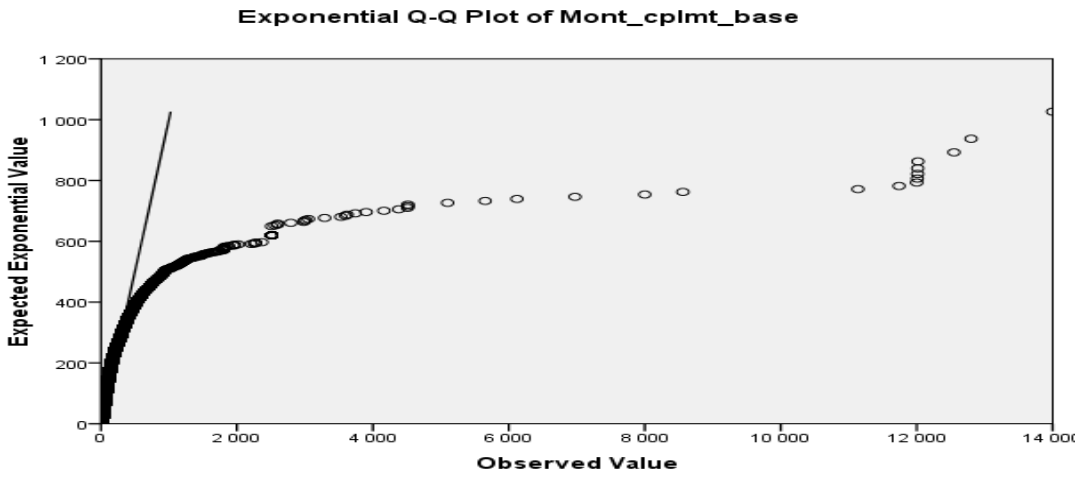

Figure 3.1 : Ajustement Exponentiel des montants de sinistres en pharmacie.

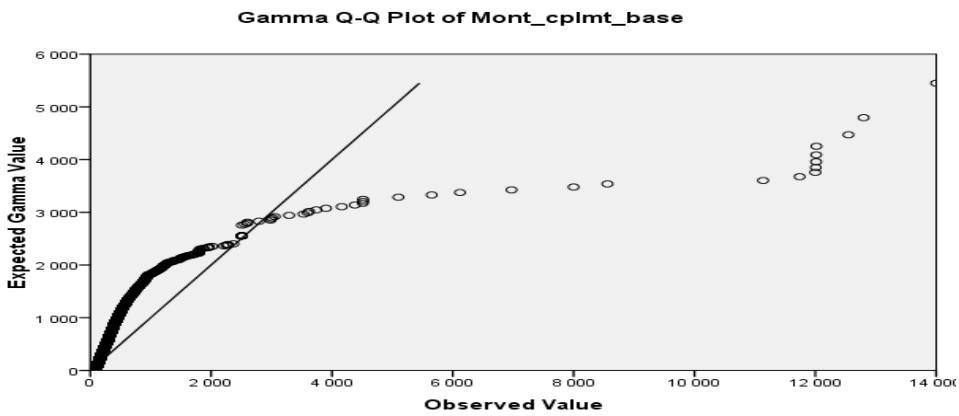

Figure 3.2 : Ajustement gamma des montants de sinistres en pharmacie.

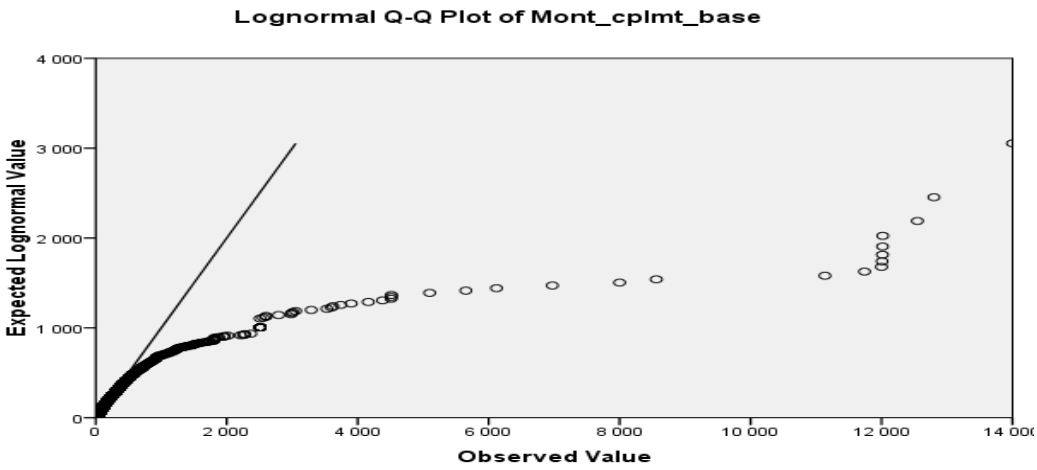

Figure 3.3 : Ajustement Lognormal des montants de sinistres en pharmacie.

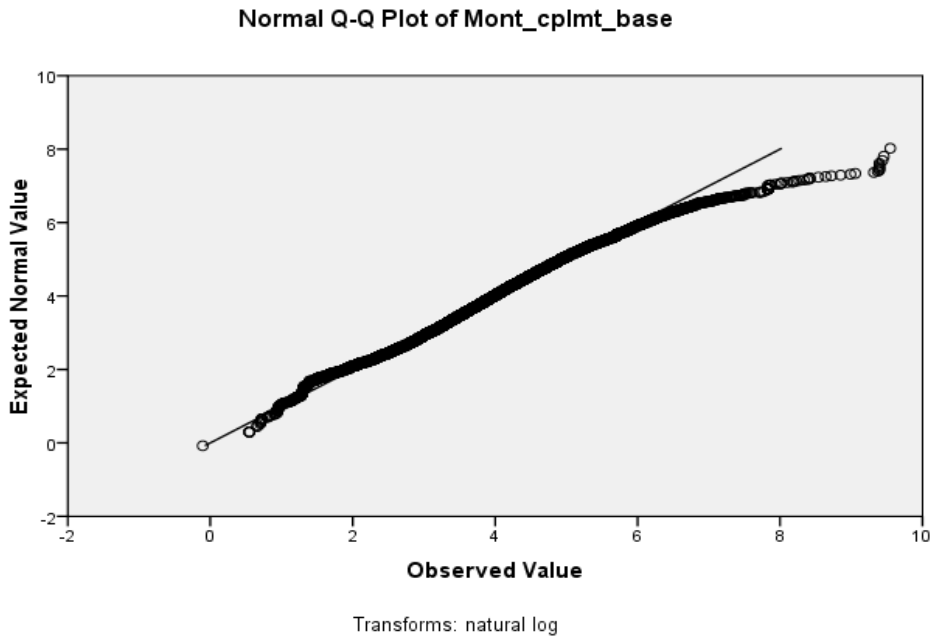

Figure 3.4 : Ajustement normal du logarithme des montants de sinistres en pharmacie.

Il apparaît donc qu'un ajustement par une loi normale de la transformation logarithmique du montant des sinistres convient pour le poste pharmacie. Le tableau ci-dessous fournit les paramètres estimés du modèle :

Estimated Distribution Parameters

		Mont_cplmt_base
Normal Distribution	μ (moyen)	**3,9733**
	σ (écart type)	**0,97537**

Tableau 3.5(bis) : Estimation des paramètres de l'ajustement en pharmacie.

Si on désigne par X la variable représentant les coûts de sinistres, alors la variable Y = ln X~Normale (μ, σ) est distribuée suivant une loi normale de moyenne μ et d'écart type σ.

Ainsi le **coût moyen d'un sinistre** en pharmacie s'écrit :

$$\boxed{E(X) = E\,(e^{Y}) = M_Y(1) = \exp\left(\mu + \frac{1}{2}\sigma^2\right)}$$

Avec $M_Y(1)$: la fonction génératrice des moments en t = 1 de la loi normale.

II.2.1.2 Récapitulation des modèles de montants de sinistres retenus pour chaque poste:

Le coût moyen des sinistres est propre à chaque poste maladie ; il est indépendant de la police assurée. Pour tous les postes, les modèles retenus correspondent à un ajustement de la transformation logarithmique des montants complémentaires de sinistre par une loi normale ou par une loi Gamma (voir annexes de 4 à 10). Les montants moyens de sinistres sont déduits par le biais de la fonction génératrice des moments correspondant respectivement à ces deux lois.

- Dans le cas où $y = \ln X \sim$ Normale (μ, σ), le coût moyen est donné par :

$$\mathbf{E(x)} = \mathbf{E}\,(\mathbf{e^Y}) = M_Y(1) = \exp\left(\mu + \frac{1}{2}\sigma^2\right).$$

- Dans le cas où $y = \ln X \sim$ Gamma (a, c), le coût moyen est donné par :

$$\mathbf{E(x) = E}\,(\mathbf{e^Y}) = M_Y(1) = \frac{1}{(1-\frac{1}{c})^a}.$$

Pour voir les détails des graphiques Q-Q plot avec les paramètres estimés pour tous les postes, le lecteur pourra se rapporter aux annexes de 4 à 10.

Nous récapitulons dans le tableau ci-dessous les modèles obtenus pour tous les postes :

Poste maladie	Modèle retenu		Coût moyen d'un sinistre E(x)
Ambulatoire	Gamma avec transformation log.	a=27.464 c=7.734	44.816 dhs
Pharmacie	Normal avec transf log.	μ=3.97 σ=0.975	85.539 dhs
Hospitalisation	Gamma avec transf log.	a=18.102 c=3.364	593.445 dhs
Analyse et radios	Gamma avec transf log.	a=21.566 c=5.017	120.789 dhs
Soins dentaires	Gamma avec transf log.	a=17.872 c=3.284	658.476 dhs
Optique	Gamma avec transf log.	a=29.151 c=5.426	379.264 dhs
Autres	Gamma avec transf log.	a=15.261 c=2.747	999.516 dhs

Tableau 3.6 : Tableau récapitulatif de la modélisation des coûts moyens de tous les postes.

II.3 Application à la détermination de la prime pure :

Afin d'illustrer concrètement cette première approche de modélisation de la prime pure, nous allons l'appliquer à trois polices aux caractéristiques de sinistralité bien distinctes comme l'indique le tableau ci-dessous :

police		taillecode		plafondcode		agecode	
	N° police	valeur	code	valeur	code	valeur	code
Police1	2500009	327	4	30000	3	18.64	2
Police2	9402397	151	2	25000	2	18.17	1
Police3	2100371	18	1	20000	1	19.44	2

Tableau 3.7 : Tableau des polices illustrant le calcul de la prime pure.

Il faut rappeler que dans cette première approche, seule la variable taillecode est retenue comme variable explicative dans la modélisation de la fréquence.

Pour chacune de ces polices, la prime pure dans un poste donné est égale au produit de sa fréquence moyenne de sinistres par le coût moyen d'un sinistre dans le poste. Cela se traduit par **la relation suivante :**

$$\textbf{Prime pure } (policei)_{poste} = \mathbf{E\ (Ni/Xi)\ .E(x)} = e^{(\hat{\beta}_0)}.\ [\ \prod_{j=1}^{5} e^{\text{taillecodej(i)}\hat{\beta}_j}\ ].\ \mathbf{E(x)}$$

La modalité taillecode_j de la variable taillecode vaut 1 si la police i y appartient et 0 sinon. La détermination de la prime pure correspondante aux trois polices fournit les résultats récapitulés dans le tableau ci-dessous :

Poste		paramètres	Fréquence moyenne.	Cout moyen	Prime pure
Ambulatoire	Police1	$\hat{\beta}_0$=7.602 ; $\hat{\beta}_4$=-1.3	546.6	44.816	24496.77 dhs
	Police2	$\hat{\beta}_0$=7.602 ; $\hat{\beta}_2$=-2.729	130.143	44.816	5832.5
	Police3	$\hat{\beta}_0$=7.602 ; $\hat{\beta}_1$=-3.753	46.05	44.816	2028.528
Pharmacie	Police1	$\hat{\beta}_0$=7.763 ; $\hat{\beta}_4$=-1.28	648.365	85.539	52894.32
	Police2	$\hat{\beta}_0$=7.763 ; $\hat{\beta}_2$=-2.61	148.125	85.539	12670.464
	Police3	$\hat{\beta}_0$=7.763 ; $\hat{\beta}_1$=-3.77	54.077	85.539	4625.692
Hospitalisation	Police1	$\hat{\beta}_0$=5.434 ; $\hat{\beta}_4$=-1.66	43.51	593.445	25820.791
	Police2	$\hat{\beta}_0$=5.434 ; $\hat{\beta}_2$=-2.47	19.46	593.445	11548.439
	Police3	$\hat{\beta}_0$=5.434 ; $\hat{\beta}_1$=-3.75	5.496	593.445	3261.573
Analyses & Radios	Police1	$\hat{\beta}_0$=6.629 ; $\hat{\beta}_4$=-1.57	155.859	120.789	18826.052
	Police2	$\hat{\beta}_0$=6.629 ; $\hat{\beta}_2$=-3.05	35.56	120.789	4295.256
	Police3	$\hat{\beta}_0$=6.629 ; $\hat{\beta}_1$=-3.91	15.132	120.789	1827.779
Soins dentaires.	Police1	$\hat{\beta}_0$=6.261 ; $\hat{\beta}_4$=-1.98	72.312	658.476	47615.716
	Police2	$\hat{\beta}_0$=6.261 ; $\hat{\beta}_2$=-3.10	23.58	658.476	15526.864
	Police3	$\hat{\beta}_0$=6.261 ; $\hat{\beta}_1$=-3.92	10.48	658.476	6900.828
Optique	Police1	$\hat{\beta}_0$=5.744 ; $\hat{\beta}_4$=-1.64	60.56	379.264	22970.8
	Police2	$\hat{\beta}_0$=5.744 ; $\hat{\beta}_2$=-2.93	16.55	379.264	6275.53
	Police3	$\hat{\beta}_0$=5.744 ; $\hat{\beta}_1$=-3.61	8.43	379.264	3196.97
Autres	Police1	$\hat{\beta}_0$=2.901 ; $\hat{\beta}_4$=-1.26	5.132	999.5165	5129.918
	Police2	$\hat{\beta}_0$=2.901 ; $\hat{\beta}_2$=-1.51	4.004	999.5165	4002.064
	Police3	$\hat{\beta}_0$=2.901 ; $\hat{\beta}_1$=-2.445	1.58	999.5165	1582.634
Prime Totale	Police1				**197754.373**
	Police2				**60151.116**
	Police3				**21624.001**

Tableau 3.8 : Application à la détermination de la prime pure selon la première approche.

Nous allons calculer les primes correspondantes à ces trois polices selon la deuxième approche de modélisation en vue d'une étude comparative.

Chapitre III: Présentation de la deuxième approche de tarification à priori.

Introduction :

Dans ce chapitre, nous proposons une deuxième méthode de tarification à priori en faisant recours à la régression logistique pour modéliser la fréquence des sinistres. L'idée consiste à estimer la prime pure dans chaque poste à partir de sa charge annuelle moyenne de sinistralité par la méthode graphique du Q-Q plot. Ainsi, la prime pure obtenue reste la même pour toutes les polices assurées. Ne faut-il pas appliquer un coefficient modérateur propre à chaque police afin de tenir compte de ses caractéristiques de sinistralité ?

La méthodologie adoptée pour estimer un tel coefficient fait recours à la régression logistique. Celle-ci permet d'expliquer la probabilité de survenance en fonction des caractéristiques (taille, âge moyen, plafond) des polices assurées.

Nous présentons explicitement la démarche avec le poste pharmacie, et les résultats obtenus pour les autres postes sont disponibles dans les annexes.

III.1 Modélisation de la fréquence :

Le modèle de régression logistique vise à décrire la liaison entre une variable binaire et un ensemble de variables explicatives qui peuvent être de nature quelconque. Ce modèle a été introduit par Day et Kerridge (1967), Cox et al. (1970). La régression logistique est un outil pour estimer la probabilité de survenance d'un événement.

Dans notre contexte, il s'agit d'expliquer la probabilité de survenance d'un sinistre maladie durant une période de couverture égale à une année par les variables comme la taille des polices, l'âge moyen et le plafond.

III.1.1 Description du modèle logistique :

A chacun des n contrats groupes du portefeuille, on associe une variable aléatoire de Bernoulli définie comme suit :

$$I_i = \begin{cases} 1 & \text{si la police i produit au moins un sinistre sur la période.} \\ 0 & \text{sinon} \end{cases}$$

On pose : la probabilité $(I_i=1) = E(I_i) = q_i \in [0 ; 1]$ pour i =1…….n.

La quantité à expliquer q_i est la probabilité que la police i produise au moins un sinistre en fonction de ses caractéristiques résumées dans un vecteur de risque $X = (X_1,\ldots\ldots X_J)$ avec $(X_1,\ldots\ldots X_J)$ est un vecteur multidimensionnel dont les composantes sont les mesures des J variables aléatoires $X_1,\ldots\ldots X_J$.

Plus exactement, la régression logistique mesure combien l'événement « causer au moins un sinistre » est plus probable que l'événement « ne pas provoquer de sinistre ».

C'est un cas particulier des modèles GLM avec comme fonction de lien le logit. Le lien logit est défini comme le logarithme népérien du rapport $\frac{q_i}{1-q_i}$. Cela conduit donc à un modèle qui lie $q_i = P(I_i=1/X=x)$ et le vecteur X des prédicteurs sous la relation non linéaire suivante :

$$\text{Logit}(q_i) = \text{Ln}\left(\frac{q_i}{1-q_i}\right) = \sum_{j=0}^{J} X_j \beta_j .$$

On en déduit le modèle logistique des probabilités q_i tel que :

$$\mathbf{q_i = P(I_i=1/x) = \frac{e^{\sum_{j=0}^{J} x_j \beta_j}}{1 + e^{\sum_{j=0}^{J} x_j \beta_j}}}$$

Où $\beta = (\beta_0, \ldots\ldots\ldots\ldots, \beta_J)$ sont les paramètres inconnus du modèle qu'il faut estimer.

Le modèle logistique est de type multiplicatif (ou log additif) : le coefficient β_j mesure en terme additif le changement dans le logarithme de la côte des risques pour chaque unité de changement dans la variable X_j ; $(1 \leq j \leq J)$.

III.1.2 Estimation des paramètres du modèle :

Pour l'estimation des coefficients d'un modèle de régression logistique, la méthode généralement utilisée est celle du maximum de vraisemblance.

Soit un échantillon de n observations indépendantes (x_i , I_i). Chaque individu est caractérisé par un vecteur multidimensionnel X= (X1,…….XJ) définissant un état. On note $\{x_m, m \in M\}$ l'ensemble des états indicés par :

$m \in M = \{M_1 \text{ x} \dots \dots \dots \text{x} M_J\}$ où M_J représente le nombre de modalités de la variable x_j. Les observations de I_i associées à l'état xm sont indépendantes et suivent une loi de Bernoulli B (p_m) où $p_m = \pi(x_m) = P(I_i=1/Xi = x_m)$.

On pose : $Z_m = \sum_{i \epsilon E_m} I_i$ avec $E_m = \{i / x_i = x_m\}$;

Z_m suit la loi binomiale de paramètres $(n(x_m), p_m)$.

La vraisemblance de l'échantillon est donc : $L(\beta, z) = \prod_{m=1}^{M} \left(p_m^{Z_m} (1 - p_m)^{n(x_m) - Z_m} \right)$

D'où : $\text{Ln } [L(\beta, z)] = \sum_{m=1}^{M} \left(Z_m \ln\left(\frac{p_m}{1 - p_m}\right) + n(x_m) \ln(1 - p_m) \right)$.

Sachant que $p_m = \frac{e^{\beta' \tilde{x}_m}}{1 + e^{\beta' \tilde{x}_m}}$

On obtient : $\text{Ln } [L(\beta,z)] = \sum_{m=1}^{M} Z_m \beta' \tilde{x}_m - \sum_{m-1}^{M} n(x_m) \ln\left(1 + e^{\beta' \tilde{x}_m}\right)$

Où $\tilde{x}_m$ est un vecteur décrivant l'état x_m

L'estimateur $\hat{\beta}$ du maximum de vraisemblance vérifie le système d'équation :

$$\left\{ \frac{\partial L(\beta)}{\partial \beta} \right\}_{\beta = \hat{\beta}} = 0$$

On obtient par dérivation de Ln [L(β,z)]:

$$\sum_{m=1}^{M} n(x_m) \tilde{x}_m^i \left[\frac{e^{\hat{\beta}' \tilde{x}_m}}{1 + e^{\hat{\beta}' \tilde{x}_m^i}} \right] = \sum_{m=1}^{M} Z_m \tilde{x}_m^i \qquad \forall \text{ i}, \quad 1 \leq i \leq J_M$$

Avec J_M : nombre de composantes du vecteur β ; il est en général différent de J (nombre de prédicteurs).

Ce système d'équations, non linéaires en β, n'a pas de solution analytique. La méthode de Newton-Raphson fournit une solution à ce système d'équations en utilisant de manière itérative une approximation linéaire.

III.1.3 Présentation du modèle pour le poste pharmacie :

Afin de construire le meilleur modèle logistique, compte tenu des données, nous utiliserons une méthode de sélection des variables dite méthode backward (descendante). Il s'agit de partir de la meilleure régression avec toutes les variables, puis à chaque pas, des tests permettent d'éliminer éventuellement des variables déjà introduites qui ne seraient plus significativement explicatives compte tenu du critère introduit dans le modèle.

Le tableau ci-dessous récapitule les modèles obtenus avec le poste pharmacie :

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	agecode	-,260	,169	2,377	1	,123	,771
	Plafondcode	-,024	,186	,017	1	,895	,976
	Taillecode	,935	,445	4,419	1	,036	2,548
	Constant	1,562	,662	5,568	1	,018	4,769
Step 2[a]	agecode	-,264	,167	2,507	1	,113	,768
	Taillecode	,918	,424	4,683	1	,030	2,505
	Constant	1,537	,634	5,879	1	,015	4,650
Step 3[a]	Taillecode	,636	,363	3,072	1	,080	1,890
	Constant	1,187	,578	4,212	1	,040	3,277

Tableau 3.9 : Régression backward du poste pharmacie.

Ainsi, le meilleur modèle retenu est celui qui retient la variable taillecode comme seule variable explicative. Toutefois, dans la colonne «sign , le test de wald indique que cette variable n'est pas significative au seuil de 5% (car 0.08>0.05).

Afin de s'assurer d'un meilleur ajustement de nos données, nous avons utilisé le test de Hosmer et Lemeshow disponible dans le logiciel SPSS.

III.1.3.1 Test de Hosmer et Lemeshow :

Hosmer et Lemeshow (1989) proposent la statistique HL qui consiste à comparer les fréquences empiriques et ajustées dans 10 classes obtenues en ordonnant et en regroupant les états $\{x_m, m \in M\}$ par décile.

- Le 1er groupe : g = 1 est composé des états tel que $\hat{\pi}_0(m) < 0{,}10$.
- Le 10ème groupe : g = 10 est relatif aux états tel que $\hat{\pi}_0(m) > 0{,}9$.

Lorsque la distance entre les fréquences observées et prédites, au moyen d'une statistique χ^2 de Pearson, est petite, on considère que le modèle est bien calibré.

La statistique de (HL) est :

$$\text{HL} = \sum_g \frac{\left(O_g - n_g \hat{\pi}_0(g)\right)^2}{n_g \hat{\pi}_0(g)(1 - \hat{\pi}_0(g))}$$

Où : $O_g = \{z_m / x_m \in g\}$, nombre observé de l'événement $I_i = 1$ dans la classe g ;

$n_g = \{n_m / x_m \in g\}$, fréquence d'appartenance à la classe g ;

$\hat{\pi}_0(g)$: la fréquence estimée sous H_0 de l'événement $I_i = 1$ dans la classe g.

Le test est le suivant : H_0 : Le modèle est adéquat ;

H_1 : Le modèle n'est pas adéquat.

Sous H_0 : HL suit approximativement une loi de χ^2 à (nombre de classes-2) degrés de liberté.

A un niveau de signification $\alpha = 5\%$, on accepte H_0 si la p-value est supérieure à 0,05.

Sinon, si la p-value est inférieure à 0,05, le modèle est rejeté.

Le test de Hosmer et lemeshow indique que notre modèle est bien calibré comme en illustre le tableau ci-dessous :

Hosmer and Lemeshow Test

modéle	Chi-square	df	Sig.
1	11,507	8	0,175
2	8,531	6	0,202
3	1,380	2	0,502

Tableau 3.10 : Test de Hosmer Lemeshow du poste Pharmacie.

On note une p-value de 0.502, supérieure à 0,05 pour le modèle retenu par la régression backward.

Si l'ajustement du modèle est adéquat, son inconvénient se situe dans la non significativité de la variable taillecode. Dans l'optique d'obtenir le meilleur modèle possible, nous avons utilisé le logiciel XLSTAT diposant d'une option « meilleur modèle » qui retient le meilleur modèle sur la base du critère spécifié. Il suffit d'indiquer le nombre minimal et le nombre maximal de variables dans le modèle.

III.1.3.2 Sélection du modèle :

En choisissant le critère de Wald, nous avons obtenu le modèle ci-dessous pour le poste pharmacie. Pour les autres postes, le lecteur pourra se reporter aux annexes de 11 à 17.

[redacted]

Tableau IV.11 : Synthèse de la sélection des variables pour le poste pharmacie.

Le test de Wald montre que toutes les variables retenues dans ce modèle sont significatives au seuil de 5%. Nous avons donc retenu les variables **agecode** et **plafondcode** pour le poste pharmacie.

III.1.3.3 Estimation des paramètres du modèle :

Pour l'estimation des coefficients d'un modèle de régression logistique, la méthode utilisée est celle du maximum de vraisemblance. Les résultats de la procédure d'estimation par le maximum de vraisemblance sont donnés dans le tableau suivant :

[redacted]

Tableau 3.11 : Estimation des paramètres du modèle logistique pour le poste pharmacie.

On note toutefois l'existence de modalités non significatives au sens de Wald notamment pour la variable plafondcode au seuil de 5%.

Cela pourrait être dû à l'insuffisance dans le fichier des prestations d'observations correspondantes à ces modalités, de nature à réellement discriminer la variable binaire expliquée.

Le modèle retenu s'écrit comme suit :

Logit (q_i) = 1.489-0,965* agecode-1-0.188*agecode-2 -2.821*agecode-3 +44.283*agecode-4-2.348* agecode-5 + 42.214* agecode-6 - 0,00* agecode-7-0,00* plafondcode-1-0.511*plafondcode-2 - 0,896*plafondcode-3 +24.633*plafondcode-4+ 24.982* plafoncode-5 – 20.927* plafondcode-6 – 2.010* plafondcode-7.

III.1.3.4 Evaluation du pouvoir discriminant du modèle logistique:

L'évaluation du pouvoir discriminant est une autre manière de tester la qualité d'ajustement du modèle de régression logistique. Elle indique la proportion des cas qui sont bien classés. Pour cela elle fait intervenir les notions de sensibilité et de spécificité.

La sensibilité est définie comme la probabilité de classer l'individu dans la catégorie y = 1 étant donné qu'il est effectivement observé dans celle-ci :

$$\text{Sensibilité} = P\,(\text{Test+} \mid y = 1)\ .$$

La spécificité est définie comme la probabilité de classer l'individu dans la catégorie y = 0 étant donné qu'il est effectivement observé dans celle-ci :

$$\text{Spécificité} = P\,(\text{Test-} \mid y = 0)\,.$$

Comme indicateur de la capacité du modèle à discriminer, **on utilisera la courbe ROC.**

La courbe ROC (Receiver Operating Characteristics) permet de visualiser la performance d'un modèle, et de la comparer à celle des autres modèles. Les termes utilisés viennent de la théorie de détection du signal. La courbe ROC est un moyen d'exprimer la relation entre la sensibilité et la spécificité d'un test diagnostique. La surface sous cette courbe nous permet d'évaluer la précision du modèle pour discriminer les outcomes positifs (y = 1) des outcomes négatifs (y = 0).

L'aire sous la courbe (ou Area Under the Curve – AUC) est un indice synthétique calculé pour les courbes ROC. L'AUC correspond à la probabilité pour qu'un événement positif ait une probabilité donnée par le modèle plus élevée qu'un événement négatif. Pour un modèle idéal, on a AUC = 1, pour un modèle aléatoire, on a AUC = 0,5. On considère habituellement que le modèle est bon dès lors que la valeur de l'AUC est supérieure à **0,7**. Un modèle bien discriminant doit avoir une AUC entre 0,87 et 0,9. Un modèle ayant une AUC supérieure à 0,9 est excellent.

On retiendra la règle de décision suivante :

Si $AUC = 0.5$, il n'y a pas de discrimination.

Si $0.7 \leq AUC \leq 0.87$, la discrimination est acceptable.

Si $0.87 \leq AUC \leq 0.9$, la discrimination est excellente.

Si $AUC \geq 0.9$, la discrimination est exceptionnelle.

Dans le cas de notre modèle, on a obtenu la courbe ROC ci-dessous :

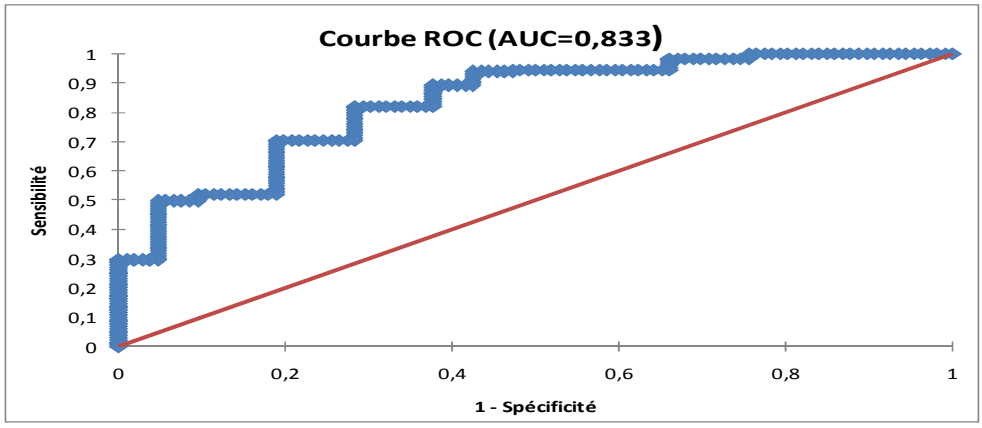

Figure 3.6: Courbe ROC du modèle logistique pour le poste pharmacie.

Avec un indice AUC égal à 0.833, on peut dire que le pouvoir discriminant du modèle est acceptable.

III.2 Modélisation de la charge annuelle de sinistralité des polices :

La démarche et les outils utilisés pour modéliser la charge sont les mêmes que dans le cas de la première approche. Seulement cette fois-ci, au lieu d'estimer le coût moyen d'un sinistre d'une police assurée pour chaque poste maladie, nous cherchons à estimer sa charge annuelle moyenne, tous sinistres confondus dans le dit poste.

III.2.1 Présentation du modèle pour le poste pharmacie :

Nous présentons ici le Q-Q plot retenu des lois suspectées qui ajuste mieux les charges complémentaires de sinistre dans le poste pharmacie. Pour les autres postes, le lecteur pourra se référer aux annexes de 11 à 17.

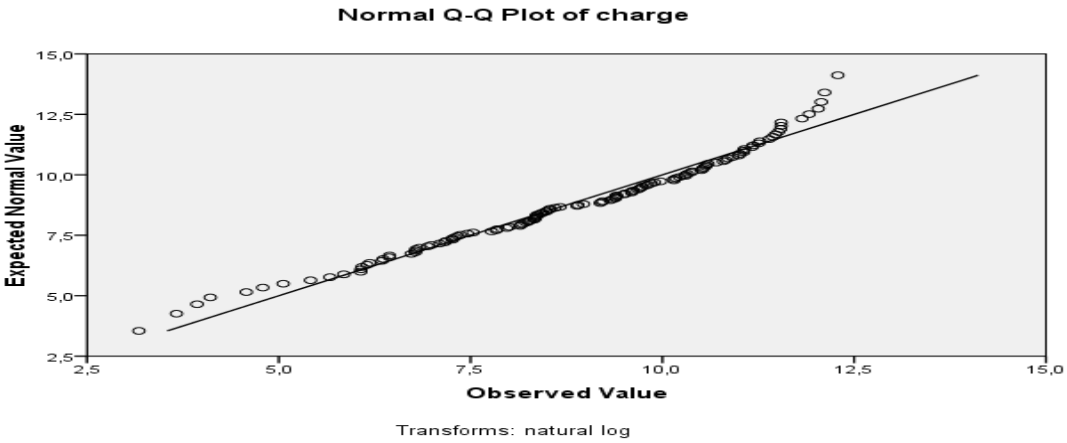

Figure3.7: Q-Q plot de la charge annuelle de sinistres pour le poste pharmacie.

Il s'est avéré qu'un ajustement par une loi normale de la transformation logarithmique des charges annuelles de sinistralité des polices convient mieux.

On peut conforter cela par l'allure quasi-normale de la courbe de la fonction densité de l'histogramme ci-dessous :

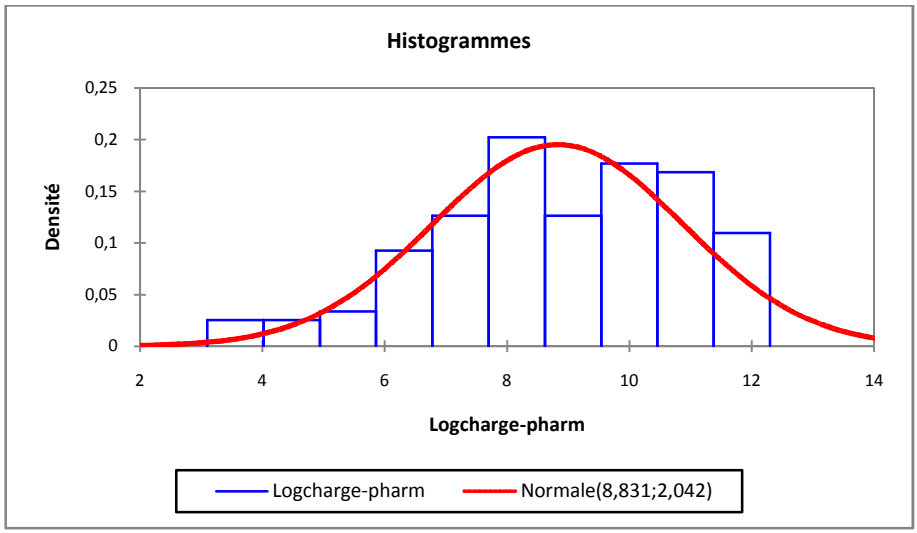

Figure 3.8 : Histogramme de la charge annuelle de sinistres pour le poste pharmacie.

Le tableau ci-dessous récapitule les paramètres estimés de la loi d'ajustement :

Paramètres estimés :	
Paramètres	**Valeurs**
μ	**8,831**
sigma	**2,042**

Tableau 3.12 : Paramètres de l'ajustement normal de la charge annuelle en pharmacie.

Afin de s'assurer d'un meilleur ajustement de la charge annuelle de sinistralité, nous avons procédé au test d'adéquation très usuel de Kolmogorov-Smirnov.

Pour comprendre le cadre théorique du test, le lecteur pourra consulter le livre de Mustapha Benyakhlef « Probabilités et Statistique Mathématique, Tome II ».

III.2.1.1 Test de Kolmogorov-Smirnov :

Nous avons choisi aléatoirement un échantillon représentatif et procédé au test suivant :

- H_0 : l'échantillon suit la loi normale.
- H_1 : l'échantillon ne suit pas la loi normale.

Les résultats du test sont donnés dans le tableau ci-dessous :

[redacted]

Tableau 3.13 : Test de Kolmogorov-Smirnov de l'ajustement normal de la charge annuelle.

Etant donné que la p-value est supérieure au seuil de signification alpha = 0.05, on peut valider l'hypothèse H_0.

L'ajustement étant adéquat, nous pouvons déterminer la charge annuelle moyenne de sinistralité avec la même formule que celle exposée dans la première approche d'estimation du coût moyen d'un sinistre.

III.2.2 Récapitulation des modèles retenus de la charge annuelle moyenne de sinistralité de chaque poste :

La charge annuelle moyenne de sinistralité est propre à chaque poste maladie ; elle est indépendante de la police assurée. Pour tous les postes, les modèles retenus correspondent à un ajustement de la transformation logarithmique des montants complémentaires de sinistre S par une loi normale.

Dans ce cas où $y = \ln S \sim \text{Normale}(\mu, \sigma)$, la charge moyenne est donnée par :

$$\mathbf{E(S) = E\,(e^{Y}) = \exp\,(\mu + \frac{1}{2}\,\sigma^{2}).}$$

Pour voir les détails des graphiques Q-Q plot avec les paramètres estimés pour tous les postes, le lecteur pourra se rapporter aux annexes de 11 à 17.

Nous récapitulons dans le tableau ci-dessous les modèles obtenus pour tous les postes :

Poste maladie	Modèle retenu		Charge annuelle moyenne E(S) (DH)
Ambulatoire	Normal avec transformation log.	μ =7.826 σ =1.914	42518.65
Pharmacie	Normal avec transf log.	μ=8.831 σ=2.042	55043.66
Hospitalisation	Normal avec transf log.	μ =8.144 σ =1.911	21374.64
Analyse et radios	normal avec transf log.	μ =8.029 σ =1.88	17965.346
Soins dentaires	normal avec transf log.	μ =9.184 σ =1.653	38185.24
Optique	Normal avec transf log.	μ =8.52 σ =1.682	20734.66
Autres	normal avec transf log.	μ =7.082 σ =1.893	7141.83

Tableau 3.14: Tableau récapitulatif de la modélisation de la charge de tous les postes.

III.3 Application à la détermination de la prime pure :

Considérons une police assurée i de charge annuelle de sinistralité S. La prime pure correspondante à cette police dans un poste maladie donné se détermine par le modèle suivant :

$$\textbf{Prime pure } (policei)_{poste} = \boldsymbol{q_i}.\mathbf{E(S)} = \frac{e^{\sum_{j=0}^{J} \beta_j x_j}}{1 + e^{\sum_{j=0}^{J} \beta_j x_j}}.\mathbf{E(S)}$$

Les J modalités sont introduites sous forme binaire selon que la police y appartient (1) ou pas (0).

Pour illustrer le calcul de la prime obtenue par cette deuxième approche de modélisation, nous allons utiliser les mêmes polices aux caractéristiques bien définies dans le tableau 3.7.

Nous récapitulons les primes pures obtenues dans le tableau ci-dessous :

Poste		Probabilité q_i	Charge moyenne E(S)	Prime pure (DH)
pharmacie	Police1	0.559	55043.66	32971.15
	Police2	0.503	55043.66	27700.72
	Police3	0.786	55043.66	43264.5
ambulatoire	Police1	0.714	42512.64	30358.309
	Police2	0.729	42512.64	30996.1
	Police3	0.65	42512.64	27637.116
Hospitalisation	Police1	0.5788	21374.64	12372.43
	Police2	0.376	21374.64	8024.64
	Police3	0.746	21374.64	15947.95
Analyses & Radios	Police1	0.619	17965.346	11123.49
	Police2	0.515	17965.346	9256.65
	Police3	0.768	17965.346	13806.814
Soins dentaires.	Police1	0.5239	38185.246	20008.366
	Police2	23.58	38185.246	15701.837
	Police3	10.48	38185.246	30497.395
Optique	Police1	0.5237	20734.66	10858.74
	Police2	0.282	20734.66	5849.45
	Police3	0.906	20734.66	18802.19
Autres	Police1	1	7141.83	7141.83
	Police2	0.724	7141.83	5171.55
	Police3	0.569	7141.83	4064.10
Prime Totale	Police1			**124834.314**
	Police2			**102700.947**
	Police3			**154020.065**

Tableau 3.15 : Application à la détermination de la prime pure pour la deuxième approche.

Chapitre IV: Analyse comparative des deux approches de tarification.

Nous essayons dans ce chapitre de dégager les avantages et les limites des deux approches de tarification présentées dans ce travail.

Il faut souligner toutefois que les résultats générés par les modèles statistiques dépendent de la fiabilité des données et de leur volume en vertu de la loi des grands nombres. A titre de comparaison, nous allons nous appuyer sur les trois polices utilisées ci-haut en guise de tarification pour illustrer succinctement les différences de primes dégagées par les deux modèles.

Polices	*Primes obtenues avec la 1ere approche (DH)*	*Primes obtenues avec la 2eme approche (DH)*
Police 1	197754.373	124834.314
Police 2	60151.116	102700.247
Police 3	21624.001	154020.65

Tableau 3.16: Tableau comparatif des primes obtenues avec les deux approches.

Cet échantillon de trois polices illustre l'écart important entre les deux approches.

La deuxième approche, basée sur la modélisation de la fréquence par une régression logistique, a tendance à surestimer la prime pure ; c'est-à-dire que les probabilités estimées n'ont pas réussi à lisser la charge annuelle moyenne de sinistralité des polices. Elles sont généralement proches de 1 ; cela se traduit par le fait que certaines modalités des variables explicatives retenues dans le modèle n'expliquent pas significativement la variable binaire modélisée, soit par leur non représentativité dans une proportion voisine dans les deux composantes (produire au moins un sinistre sur un an ou pas) de la variable expliquée.

Le fichier des prestations n'est pas assez exhaustif pour booster cette régression logistique. L'idée qu'en risque maladie qu'une police d'un contrat groupe ne produise aucun sinistre sur un an peut paraître difficilement réalisable.

Globalement, cette deuxième approche n'a pas donné les résultats escomptés puisqu'elle est beaucoup plus en deçà des tarifs standards de la tarification nationale de référence.

Faire la tarification avec ce modèle pourrait induire des risques non concurrentiels dans le marché marocain de l'assurance maladie complémentaire.

Par contre, la première approche basée sur une modélisation GLM de la fréquence des sinistres par une binomiale négative, est beaucoup plus proche de la tarification nationale de référence. Elle est très pratique puisqu'elle retient comme seule variable explicative la taille des groupes, information qui soit la plus accessible possible. Faut il souligner que cette prime pure reste une référence dans la négociation commerciale des contrats ; elle traduit simplement la capacité de sinistralité des polices. La compagnie pourrait l'appliquer des coefficients multiplicatifs à sa guise afin de l'ajuster à la concurrence dans le marché.

Afin d'automatiser la tarification par cette première approche, nous avons conçu une application sur Access qui fait l'objet de la quatrième de ce rapport.

Partie IV

Application de Tarification des contrats groupe.

Le but de cette partie consiste à présenter sommairement l'application informatique que nous avons conçue afin d'automatiser le fonctionnement des modèles élaborés et retenus dans ce travail.

La conception de l'application requiert l'utilisation des notions sur les bases de données relationnelles et un logiciel SGBD. Nous avons recouru à Access comme SGBD[3] qui dispose d'une interface conviviale.

La tâche essentielle de ce travail est de sélectionner les données nécessaires pour le calcul de la prime, à partir de la base de données initiale et des résultats des modèles générés par les logiciels statistiques (SPSS et XLSTAT) en vue d'élaborer un modèle conceptuel de données.

Ensuite, il faut implémenter ce modèle avec les objets d'ACCESS.

[3] Système de gestion des bases de données.

Chapitre I : Le modèle conceptuel des données :

Le modèle élaboré s'appuie théoriquement sur la méthode de Merise4 ; il établit des relations entre différentes entités représentées par des tables sur Access. Chaque table est caractérisée par un ensemble d'attributs qui sont les informations qu'elle contient ;ces attributs de type numérique ou texte, sont reliés entre eux selon des règles validant les formes de normalisation des bases de données relationnelles. Nous ne pouvons présenter ici toutes les notions relatives aux bases de données relationnelles, mais nous avons validé le maximum de règles pour une intégrité référentielle et une absence de redondance au sein du modèle conceptuel. Au sein d'une table, il existe un ou deux attributs clés dont la connaissance permet de déterminer tous les autres.

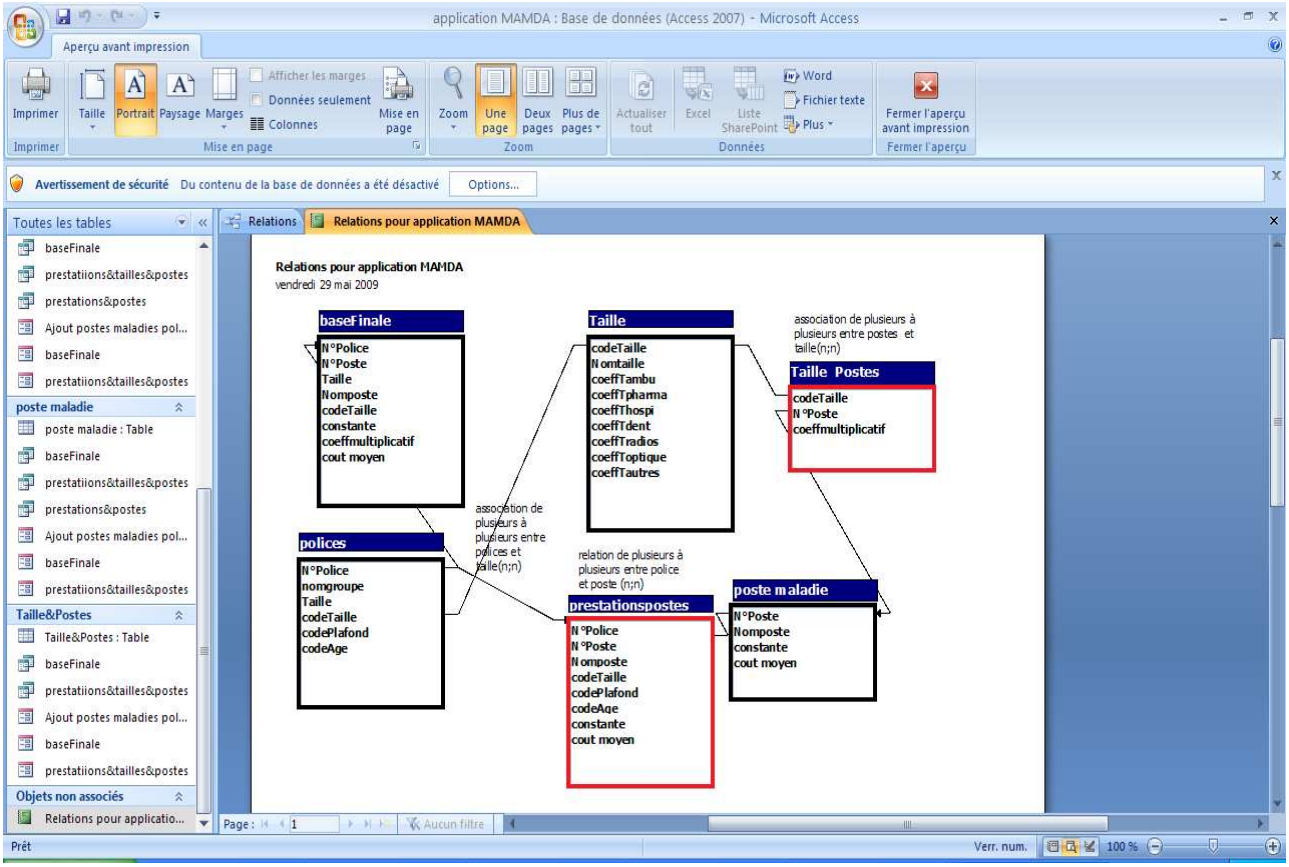

Figure 4.1 : le modèle conceptuel des données (MCD).

[4] Merise est une méthode de conception d'un systéme d'information.

Le modèle conceptuel des données est constitué à partir des objets suivants :

- **La table police** contient les informations sur les polices et leurs caractéristiques de sinistralité (codeTaille,codePlafond,codeAge) et chaque enregistrement est identifié par le champ n° police.
- **La table poste maladie** contient tous les postes avec les informations qui leur sont spécifiques comme le nom du poste , la constante du modèle et le coût moyen d'un sinistre. Chaque enregistrement est identifié par le champ n°poste.
- **La table Taille** contient les modalités de la variable taille, retenue comme seule variable explicative dans la premiere approche et les coefficients de ces modalités obtenues dans tous les postes maladie. Chaque enregistrement est identifié par le champ codeTaille.
- **La table prestations postes** est une association de plusieurs à plusieurs entre les tables polices et postes.
- **La table Taille Postes** est une association de plusieurs à plusieurs entre les tables Taille et postes maladie.
- **La table baseFinale** est une requete SQL faisant la jointure entre les tables polices , poste maladie et Taille postes sur les champs **N° police** et **N° poste**. C'est elle qui permet de calculer les fréquences moyennes de sinistralité des polices et leurs primes pures dans chaque poste. Elle a permis d'élaborer l'interface principale de l'application que nous décrirons succintement.

Chapitre II : L'interface principale de l'application :

Access dispose d'un objet très convivial appelé formulaire. Les formulaires sont des fiches liées à des tables permettant à l'utilisateur de mettre à jour les données dans les tables (ajout, suppression, modification). Il est possible d'y insérer des contrôles comme des zones de texte, des boutons de commandes ou des sous-formulaires. Les boutons de commande sont des macros qu'on peut programmer, qui exécutent des tâches au sein de l'application telle que la mise à jour ou l'accessibilité entre les objets (lancement de requêtes et ouverture de formulaires par exemple).

Nous avons ainsi créé un formulaire principal à partir de la table police et un sous-formulaire à partir de la table BaseFinale . Les deux sont joints dans l'interface principale sur le champ n° police.

Il suffit que l'utilisateur inscrit le numero d'une police dans le formulaire , il obtient les primes correspondantes à la police calculées dans tous les postes. On peut également voir la prime totale et la prime totale rapportée au nombre d'ahérents de la police.

Ensuite, un panel de commandes programmées lui permet de mettre à jour l'application comme l'ajout des contrats ou des prestations concernées par la police

Cliquer sur le champ N°police puis sur le bouton recherche police et entrer le numéro de la police à tarifer. Ne pas mettre le 0 pour les numéros commençant par 0.

Panel de commandes pour la mise à jour des données.

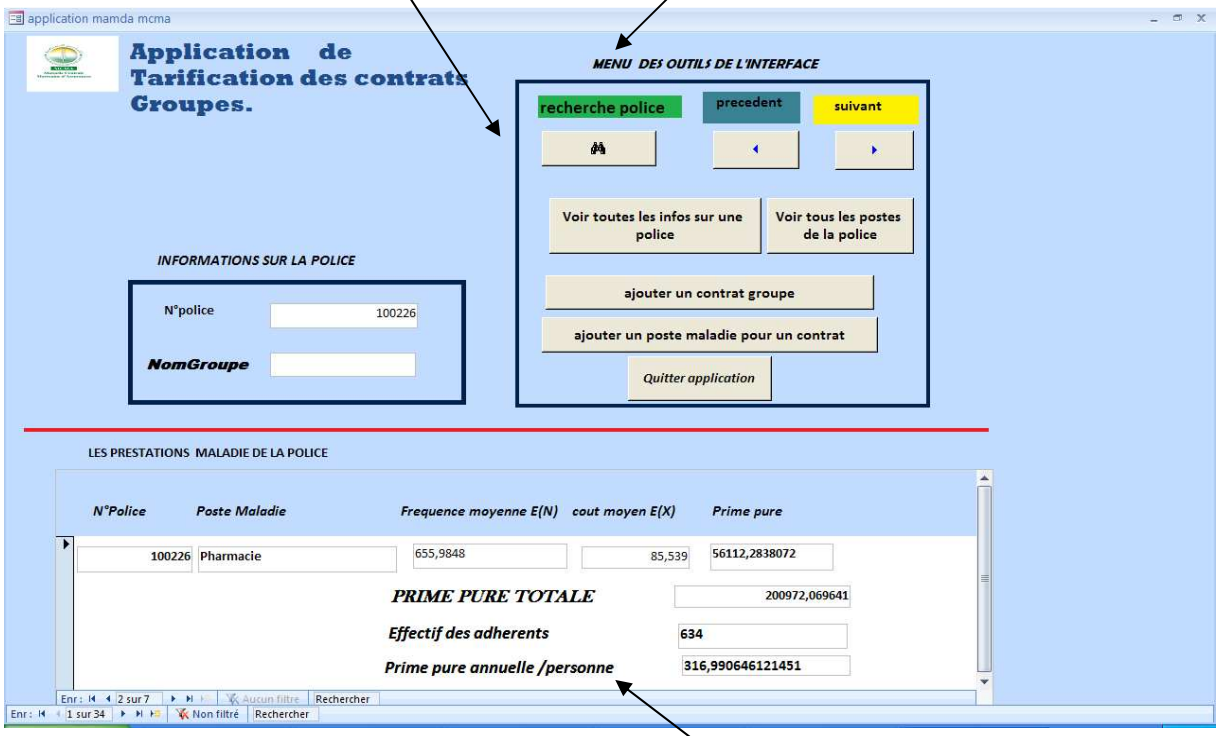

Figure 4.2 : Interface principale de l'application .

Sous -formulaire affichant les primes de tous les postes maladie pour la police. IL suffit de cliquer sur la barre de défilement de 1 à 7.

Conclusion

L'essentiel de ce travail a été consacré à l'élaboration d'une tarification à priori du ticket modérateur des soins de santé engagés par les assurés de portefeuille de la compagnie MAMDA-MCMA. Cette tarification s'est effectuée selon les deux approches, elle a pu mettre en évidence les différences enregistrées au niveau des résultats obtenus. D'abord dans la première approche, s'agissant de la modélisation, pour chaque poste de maladie, de la fréquence moyenne de sinistres, seule la variable taille du groupe, parmi les trois variables tarifaires (taille et âge moyen du groupe et plafond de remboursement), a été retenue comme variable explicative par la loi binomiale négative. Celle-ci ajuste mieux, selon le modèle GLM, cette fréquence. Alors que l'estimation du coût moyen annuel de sinistres a été obtenue, pour chaque poste, soit par un ajustement de lois normales, soit de lois gamma appliquées toutes les deux à une transformation- Log.

Dans la deuxième approche, la modélisation de la fréquence moyenne de sinistre, a été obtenue par une estimation de la probabilité de survenance d'un sinistre par une régression logistique. On a ainsi essayé d'expliquer cette probabilité en fonction des variables tarifaires susmentionnées en procédant par la méthode descendante « Backward », qui a montré que le nombre de variables retenues varie selon le poste de maladie. Quant à l'estimation du coût annuel moyen d'une police, seule une distribution normale appliquée à une transformation-Log, l'ajuste mieux.

L'analyse comparative des résultats issus des deux approches de tarification, permet de conclure que si la première approche a pu générer pour chaque poste, des primes qui s'approchent généralement du ticket modérateur moyen de chaque groupe d'assurés, au contraire la deuxième approche ,basée sur la modélisation de la fréquence des sinistres par une régression logistique, a tendance à surestimer cette prime ; les probabilités estimées n'ont pas réussi à lisser la charge annuelle moyenne de sinistralité des polices.

On peut conclure donc, que cette deuxième approche n'a pas donné les résultats escomptés puisqu'elle est beaucoup plus en deçà des tarifs standards de la tarification nationale de référence. Ce qui justifie notre choix de la première approche pour concevoir une application informatique sur Access, et qui a fait l'objet de la dernière partie de ce travail.

Aujourd'hui, la complémentaire santé s'annonce comme un nouveau produit suscitant un intérêt particulier pour les compagnies d'assurance marocaines depuis l'apparition de

l'AMO. D'où l'importance de poursuivre la recherche et le développement de modèles de tarification de ce type de prestation, afin de mieux l'adapter au marché marocain de l'assurance santé.

BIBLIOGRAPHIE :

Ouvrages :

- Artur Charpentier « Mathématiques de l'assurance Non Vie ; Tome II : Tarification et Provisionnement».
- Christian Partrat « Modélisation et simulation actuarielles ».
- EL Mustapha Benyaklef « Probabilités et Statistique mathématique, TOME II ».
- Jean claude guizouarn « Assurance Santé, Segmentation et compétitivité» ; Economica 2004.

Cours :

- Assurance Non vie, Mr Lionel Corre.
- Base de données relationnelles et Merise, Mr Khalid El Hachimi.
- Statistique Non paramétrique, Mr Layachi Baghagha.
- Théorie du risque, Mr Kamal Benchekroun.

Rapports et PFE :

- Elaboration d'un tarif individuel en assurance Maladie de Base complémentaire, Azouzi et Roland 2008.
- L'économiste du Mardi 12/4/2009.
- Le Matin lundi 27 Octobre 2008.
- Rapport annuel d'activité de la MAMDA/MCMA 2007.
- Rapport de l'Agence Nationale de l'Assurance Maladie (ANAM) édition octobre 2006.

- Tarification Automobile à priori, segmentation et modélisation des portefeuilles TPV et Frontière, HIND KAISSI 2002.

Sites Internet :

- www.assurancemaladie.ma.
- www.leconomiste.com.
- www.lematin.com.

Annexes

Annexe 1 :

Pour les actifs		Pour les pensionnés
Marins pêcheurs	**Autres secteurs**	
Le taux de cotisation due par les marins pêcheurs à la part est fixé à : ▪ **1,2%** du montant du produit brut de la vente de poisson pêché sur les chalutiers ; ▪ **1,5%** du montant du produit brut de la vente de poisson pêché sur les sardines et les palangriers. (Articles premier et deuxième du décret n° 2-05-734)	Le taux de cotisation due à la CNSS au titre de régime de l'AMO de base est fixé, pour les affilies bénéficiant de cette couverture, à : ▪ **4%** de l'ensemble des rémunérations reparti à raison de **50%** à la charge de l'employeur et **50%** à la charge du salarié. ▪ Ce taux de cotisation est majoré de **1%** de l'ensemble des rémunérations brutes mensuelles du salarié, cette majoration, due par l'ensemble des employeurs assujettis au régime de sécurité sociale, est à leur charge exclusive.	Le taux de cotisation due par les titulaires de pensions est fixé à : ▪ **4%** sur le mentant globale des pensions de base servies à condition que ce montant soit supérieur ou égal **910 DH** par mois dans le secteur de l'agriculture et à **1289 DH** par mois dans les autres secteurs. La cotisation mensuelle due par les personnes bénéficiant d'une assurance volontaire conformément à l'article 5 du dahir portant **loi n°1-72-184** est fixé à : ▪ **4%** du montant de rémunération mensuelle ayant servie de base au calcul de la dernière cotisation obligatoire au titre de la dite assurance. (Articles 3 et 4 du décret n° 2-05-734)

Tableau 1.2 : Taux de cotisation de secteur privé.

Source : Agence Nationale de l'Assurance Maladie ; Edt : octobre 2006.

Annexe 2 :

Prestations couvertes	Taux de couverture
Les groupes de prestations sont, comme désigne l'article premier et l'article 2 du décret n° 2- 05-736 :	Les taux de couverture des prestations sont fixés par groupe de celles-ci comme désigne l'article premier et l'article 2 du décret n° 2-05-736 :
***Groupe I* : Actes de médecine générale et de spécialités médicales et chirurgicales ; Actes paramédicaux ; Actes de rééducation fonctionnelle et de kinésithérapie délivrés à titre ambulatoire hors médicaments.**	**80%** de la tarification nationale de référence.
***Groupe II* : Soins liés à l'hospitalisation et aux interventions chirurgicales y compris les actes de chirurgie réparatrice et le sang et ses dérivés labiles.**	**90%** de la tarification nationale de référence. Ce taux est porté à **100%** lorsque les prestations sont rendues dans les hôpitaux publics, les établissements publics de santé et de services sanitaires relevant de l'Etat.
***Groupe III* : Médicaments admis au remboursement**	**70%** du prix public Maroc
***Groupe I V* : Lunetterie médicale, Dispositifs médicaux et implants nécessaires aux actes médicaux et chirurgicaux.**	**Forfaits** fixés dans la tarification nationale de référence.
***Groupe V* : Appareils de prothèse et d'orthèse médicales admis au remb**	**Forfaits** fixés dans la tarification nationale de référence.

Tableau 1.3 : Taux de couverture de secteur public.

Source : Agence Nationale de l'Assurance Maladie ; Edt : Octobre 2006.

Annexe 3 :

Pour les Actifs	Pour les Retraités
Conformément à l'article premier du décret n° 2-05-735(18 juillet 2005) :	Conformément à l'article 2 du décret n° 2-05-735(18 juillet 2005) :
Le taux de cotisation due à la CNOPS au titre du régime de l'assurance maladie obligatoire de base est fixé à **5%** de l'ensemble des rémunérations visées à l'article 106 de la loi 65-00, réparti à raison de **50%** à la charge de l'employeur et **50%** à la charge du salarié.	Le taux de cotisation due par les titulaires de pensions est fixé à **2,5%** du montant global des pensions de base servies dans la limite d'un montant mensuel de **70 DH** et d'un plafond annuel de **400 DH.**

Tableau 1.4 : Taux de cotisation de secteur public.

Source : Agence Nationale de l'assurance Maladie ; Edt : octobre 2006.

Annexe 4 :

Code Acte	Libellé acte	Code Acte	Libellé Acte
0	PHARMACIE CHIRURGICALE	40	ELECTRO-ENCEPHALOGRAMME
1	PHARMACIE	41	ANG
2	CONSULTATION GENERALISTE	42	GESTE COMMERCIAL
3	CONSULTATION SPECIALISTE	43	FOND D'OEIL
4	CONSULTATION PROFESSEUR	44	INFILTRATION
5	VISITE GENERALISTE	45	LASER
6	VISITE SPECIALISTE	48	IRM 1 REGION
7	ECHOGRAPHIE	49	LITHOTRIPSIE
8	PC OU AMI	50	LIT D'ACCOMPAGNANT
9	AMI OU PC CROISSANT ROUGE	51	SOINS MEDICAUX
10	ANALYSES SANS PRELEVEMENT	52	VISITE NUIT/JOUR FERIE
11	ANALYSES AVEC PRELEV. SAN	53	VISITE PSY/PROFESSEUR
13	RADIOS	54	MONITORING
14	ELECTRO CARDIOGRAMME	57	DIALYSE
15	SOINS DENTAIRES	59	REANIMATION LOURDE
16	PROTHESE DENTAIRE	62	CONSULTATION GENERALISTE
17	CHIRURGIE	63	CONSULTATION SPECIALISTE
18	SALLE D'OPERATION	65	PHARMACIE GROSSESSE
19	ANESTHESIE	66	ANALYSE SANS PRELEVEMENT
20	PERFUSION	69	ECHOGRAPHIE GROSSESSE
21	AIDE OPERATOIRE	70	TOMODENSITOMETRIE
22	REANIMATION	71	ANGIO-SCANNER
23	HOSPITALISATION	72	OSTEODENSITOMETRIE
24	TRAITEMENTS SPECIAUX	73	IRM 2 REGIONS
25	LUNETTES-MONTURE	74	ORTHOPEDIE DENTHOFACIALE
26	LUNETTES-VERRES	76	K30
27	NAISSANCE SIMPLE	77	K35
28	NAISSANCE GEMELLAIRE	78	K40
29	DIVERS	79	K45
30	CONTRE VISITE	81	COLONOSCOPIE
31	SCANNER	84	CORONAROGRAPHIE
32	HONORAIRES MEDECINS	85	ACTE COMPLEMENTAIRE A 20
33	TEST ALLERGOLOGIQUE	86	SOINS MEDICAUX A 100%
34	CHIRURGIE NON GRAVE	87	CONTROL MEDICAL
36	SALLE ACCOUCHEMENT	99	TAXES
38	VISITE PROFESSEUR		
39	FIBROSCOPIE		

Tableau 2.1 (bis) : Actes de Santés et codes correspondants.

Annexe 5 :

Correlations

			MNTPLF	taille	agemoyen
Kendall's tau_b	MNTPLF	Correlation Coefficient	1,000	,398**	,440**
		Sig. (2-tailed)	.	,000	,000
		N	103899	103899	103899
	taille	Correlation Coefficient	,398**	1,000	,279**
		Sig. (2-tailed)	,000	.	,000
		N	103899	103899	103899
	agemoyen	Correlation Coefficient	,440**	,279**	1,000
		Sig. (2-tailed)	,000	,000	.
		N	103899	103899	103899
Spearman's rho	MNTPLF	Correlation Coefficient	1,000	,470**	,457**
		Sig. (2-tailed)	.	,000	,000
		N	103899	103899	103899
	taille	Correlation Coefficient	,470**	1,000	,361**
		Sig. (2-tailed)	,000	.	,000
		N	103899	103899	103899
	agemoyen	Correlation Coefficient	,457**	,361**	1,000
		Sig. (2-tailed)	,000	,000	.
		N	103899	103899	103899

**. Correlation is significant at the 0.01 level (2-tailed).

Tableau 2.5 : Assorti d'analyse de corrélation de Kendall et de Spearman.

Annexe 6 :

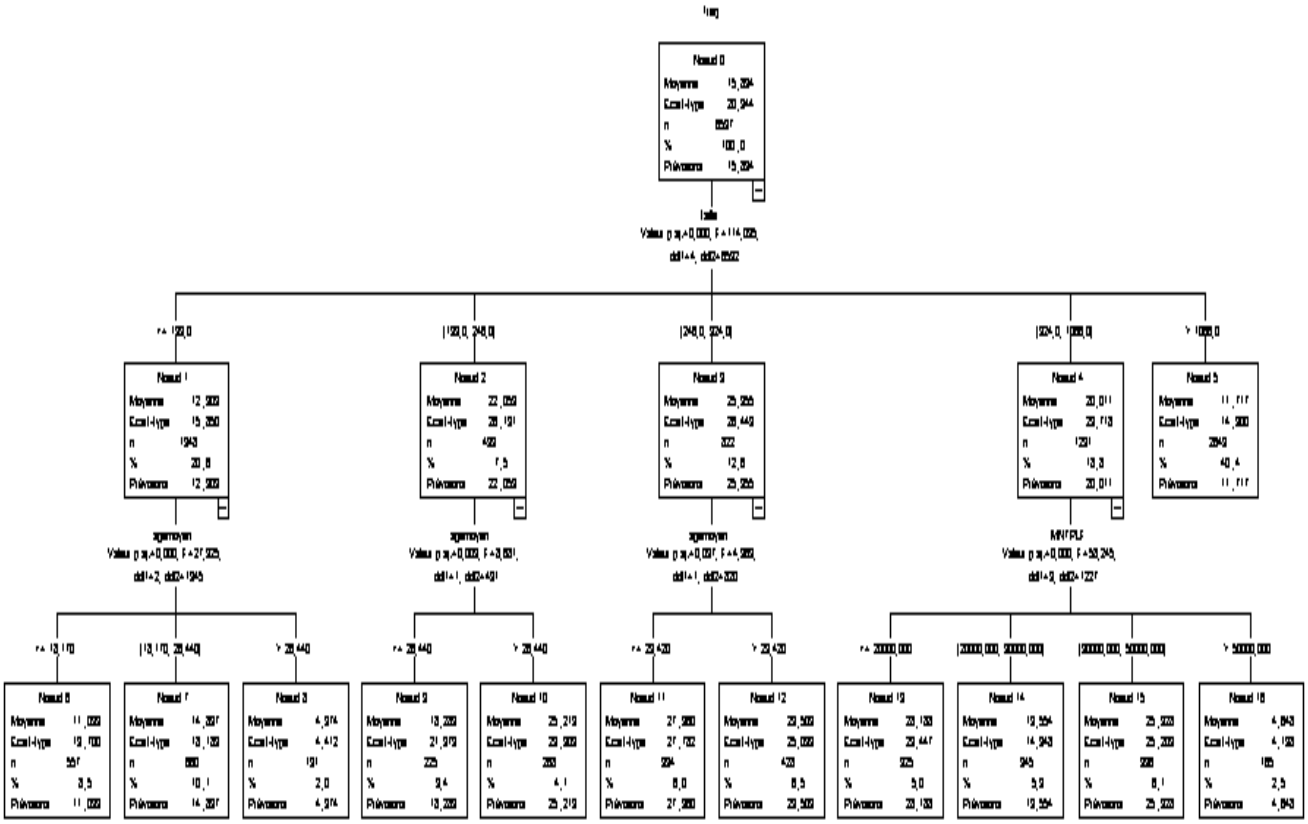

Figure 2.4 : Arbre de segmentation jusqu'au 3éme niveau pour le ticket modérateur a en prenant la fréquence de sinistres comme critère de segmentation.

Annexe 7 : Résultats de la première approche pour le poste ambulatoire.

Corrélations

		Taillecode	agecode	PLafondcode
Taillecode	Pearson Corrélation	1,000	,542**	,587**
	Sig. (2-tailed)		,000	,000
	N	130,000	130	130
agecode	Pearson Corrélation	,542**	1,000	,440**
	Sig. (2-tailed)	,000		,000
	N	130	130,000	130
PLafondcode	Pearson Corrélation	,587**	,440**	1,000
	Sig. (2-tailed)	,000	,000	
	N	130	130	130,000

Tableau 3.17: Analyse de la corrélation de Pearson pour le poste ambulatoire.

	Type III		
Source	Wald Chi-Square	df	Sig.
(Intercept)	2045,210	1	,000
Taillecode	165,894	4	,000

Tableau 3.18 : Analyse de type III de la variable Taillecode pour le poste ambulatoire.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	7,602	,4648	6,691	8,513	267,526	1	,000	2002,200	805,171	4978,822
[Taillecode_first=1]	-3,753	,4817	-4,697	-2,809	60,715	1	,000	,023	,009	,060
[Taillecode_first=2]	-2,729	,5200	-3,749	-1,710	27,551	1	,000	,065	,024	,181
[Taillecode_first=3]	-1,339	,5197	-2,357	-,320	6,635	1	,010	,262	,095	,726
[Taillecode_first=4]	-1,300	,5326	-2,344	-,256	5,956	1	,015	,273	,096	,774
[Taillecode_first=5]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1,080[b]									
(Negative binomial)	1									

Dependent Variable: frequence
Model: (Intercept), Taillecode_first

a. Set to zero because this parameter is redundant.

b. Computed based on the deviance.

Tableau 3.19 : Estimation des paramètres dans la modélisation de la fréquence des sinistres pour le poste ambulatoire.

Estimated Distribution Parameters

		Mont_cplmt_base
Gamma Distribution	Shape	27,464
	Scale	7,734

Tableau 3.20 : Estimation des paramètres dans la modélisation du cout des sinistres pour le poste ambulatoire.

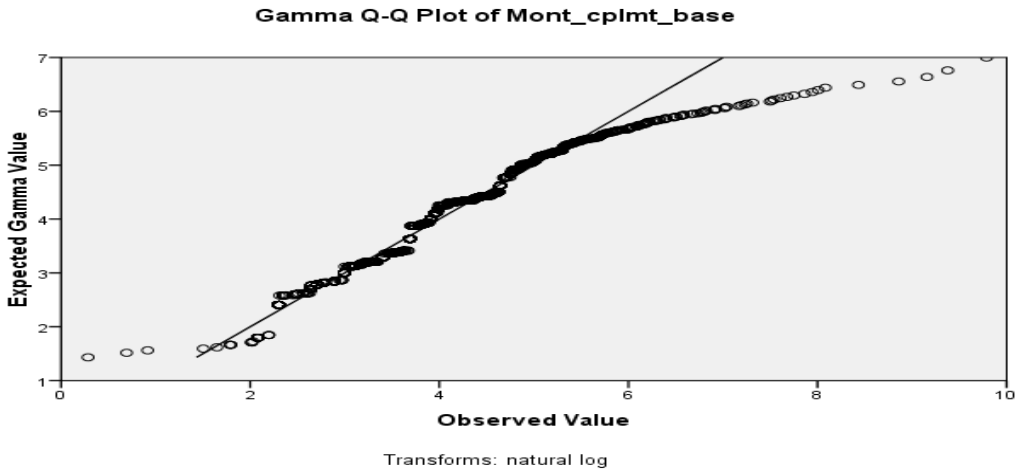

Figure 3.8 : Ajustement Gamma de la transformation logarithmique des montants de sinistre.

Annexe 8 : Résultats de la première approche pour le poste pharmacie.

Corrélations

		agecode	Taillecode	PLafondcode
agecode	Pearson Corrélation	1,000	,543**	,441**
	Sig. (2-tailed)		,000	,000
	N	129,000	129	129
Taillecode	Pearson Correlation	,543**	1,000	,586**
	Sig. (2-tailed)	,000		,000
	N	129	129,000	129
PLafondcode	Pearson Correlation	,441**	,586**	1,000
	Sig. (2-tailed)	,000	,000	
	N	129	129	129,000

Tableau 3.21 : Analyse de la corrélation de Pearson pour le poste pharmacie.

Goodness of Fit[b]

	Value	df	Value/df
Deviance	130,974	124	1,056
Scaled Deviance	130,974	124	
Pearson Chi-Square	135,327	124	1,091
Scaled Pearson Chi-Square	135,327	124	
Log Likelihood[a]	-772,430		
Akaike's Information Criterion (AIC)	1554,860		
Finite Sample Corrected AIC (AICC)	1555,348		
Bayesian Information Criterion (BIC)	1569,159		
Consistent AIC (CAIC)	1574,159		

Dépendent Variable: Fréquence

Model: (Intercept), Taillecode

Tableau 3.22: Indicateurs d'ajustement du modèle de la fréquence pour le poste pharmacie.

Annexe 9 : Résultats de la première approche pour le poste hospitalisation.

Corrélations

		agecode	PLafondcode	Taillecode
agecode	Pearson Correlation	1,000	,429**	,570**
	Sig. (2-tailed)		,000	,000
	N	107,000	107	107
PLafondcode	Pearson Correlation	,429**	1,000	,603**
	Sig. (2-tailed)	,000		,000
	N	107	107,000	107
Taillecode	Pearson Correlation	,570**	,603**	1,000
	Sig. (2-tailed)	,000	,000	
	N	107	107	107,000

Tableau 3.23: Analyse de la corrélation de Pearson pour le poste Hospitalisation.

Goodness of Fit[b]

	Value	df	Value/df
Deviance	59,296	102	,581
Scaled Deviance	59,296	102	
Pearson Chi-Square	79,202	102	,776
Scaled Pearson Chi-Square	79,202	102	
Log Likelihood[a]	-411,054		
Akaike's Information Criterion (AIC)	832,109		
Finite Sample Corrected AIC (AICC)	832,703		
Bayesian Information Criterion (BIC)	845,473		
Consistent AIC (CAIC)	850,473		

Dependent Variable: frequence
Model: (Intercept), Taillecode

Tableau 3.24 : Indicateurs d'ajustement de la fréquence du poste Hospitalisation.

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	786,657	1	,000
Taillecode	121,610	4	,000

Tableau 3.25 : Analyse de type III de la variable Taillecode du poste Hospitalisation.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	5,434	,4482	4,555	6,312	146,985	1	,000	229,000	95,134	551,232
[Taillecode=1]	-3,748	,4749	-4,679	-2,817	62,279	1	,000	,024	,009	,060
[Taillecode=2]	-2,469	,5092	-3,467	-1,471	23,510	1	,000	,085	,031	,230
[Taillecode=3]	-1,573	,5019	-2,557	-,589	9,822	1	,002	,207	,078	,555
[Taillecode=4]	-1,662	,5146	-2,671	-,654	10,436	1	,001	,190	,069	,520
[Taillecode=5]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									
(Negative binomial)	1									

Dependent Variable: frequence
Model: (Intercept), Taillecode

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Tableau 3.26 : Estimation des paramètres du modèle de la fréquence du poste Hospitalisation.

Estimated Distribution Parameters

		Mont_cplmt_base
Gamma Distribution	**Shape**	**18,102**
	Scale	**3,364**

Tableau 3.27: Estimation des paramètres dans la modélisation du cout des sinistres pour le poste Hospitalisation.

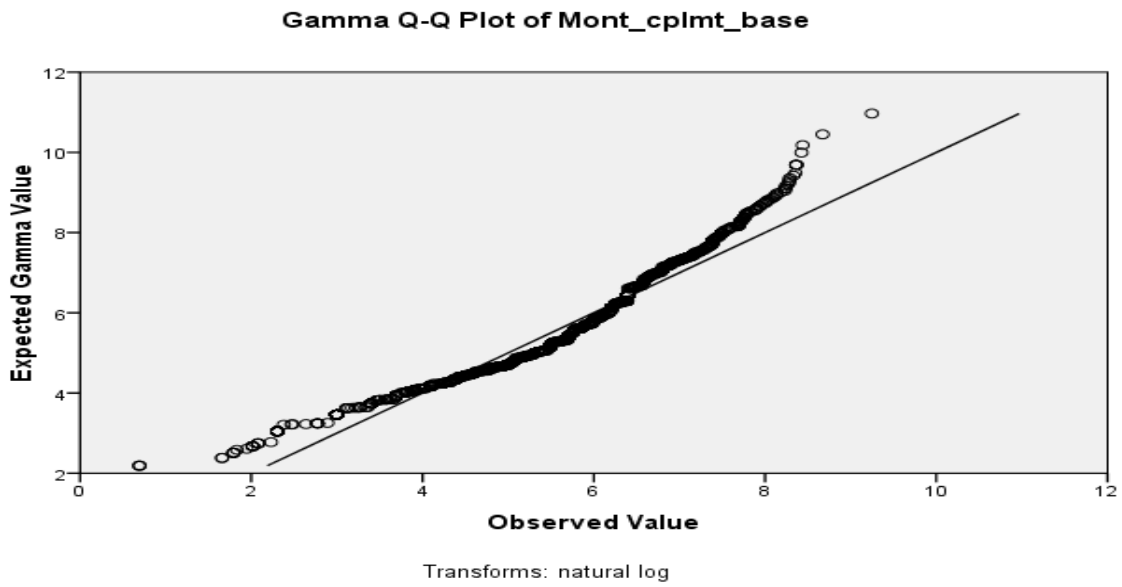

Figure 3.9: Ajustement Gamma de la transformation logarithmique des montants de sinistre.

Annexe 10 : Résultats de la première approche pour le poste Analyses et Radios.

Corrélations

		Taillecode	agecode	PLafondcode
Taillecode	Pearson Corrélation	1,000	,568**	,614**
	Sig. (2-tailed)		,000	,000
	N	124,000	124	124
agecode	Pearson Corrélation	,568**	1,000	,436**
	Sig. (2-tailed)	,000		,000
	N	124	124,000	124
PLafondcode	Pearson Corrélation	,614**	,436**	1,000
	Sig. (2-tailed)	,000	,000	
	N	124	124	124,000

Tableau 3.28: Analyse de la corrélation de Pearson pour le poste Analyses et radios.

Goodness of Fit[b]

	Value	df	Value/df
Deviance	92,295	119	,776
Scaled Deviance	92,295	119	
Pearson Chi-Square	101,722	119	,855
Scaled Pearson Chi-Square	101,722	119	
Log Likelihood[a]	-591,894		
Akaike's Information Criterion (AIC)	1193,788		
Finite Sample Corrected AIC (AICC)	1194,297		
Bayesian Information Criterion (BIC)	1207,890		
Consistent AIC (CAIC)	1212,890		

Dependent Variable: Fréquence

Model: (Intercept), Taillecode

Tableau 3.29 : Indicateurs d'ajustement de la fréquence du poste Analyses et radios.

Tests of Model Effects

	Type III		
Source	Wald Chi-Square	df	Sig.
(Intercept)	1436,435	1	,000
Taillecode	188,591	4	,000

Tableau 3.30 : Analyse de type III de la variable Taillecode du poste Radios.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	6,629	,4475	5,752	7,506	219,417	1	,000	756,600	314,736	1818,807
[Taillecode=1]	-3,912	,4660	-4,826	-2,999	70,476	1	,000	,020	,008	,050
[Taillecode=2]	-3,048	,5017	-4,031	-2,065	36,919	1	,000	,047	,018	,127
[Taillecode=3]	-1,190	,5005	-2,171	-,209	5,654	1	,017	,304	,114	,811
[Taillecode=4]	-1,578	,5130	-2,584	-,573	9,464	1	,002	,206	,075	,564
[Taillecode=5]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									
(Negative binomial)	1									

Dependent Variable: Fréquence
Model: (Intercept), Taillecode

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Tableau 3.31. : Estimation des paramètres du modèle de la fréquence du poste Radios.

Estimated Distribution Parameters

		Mont_cplmt_base
Gamma Distribution	**Shape**	**21,566**
	Scale	**5,017**

Tableau 3.32: Estimation des paramètres dans la modélisation du coût des sinistres pour le poste Radios.

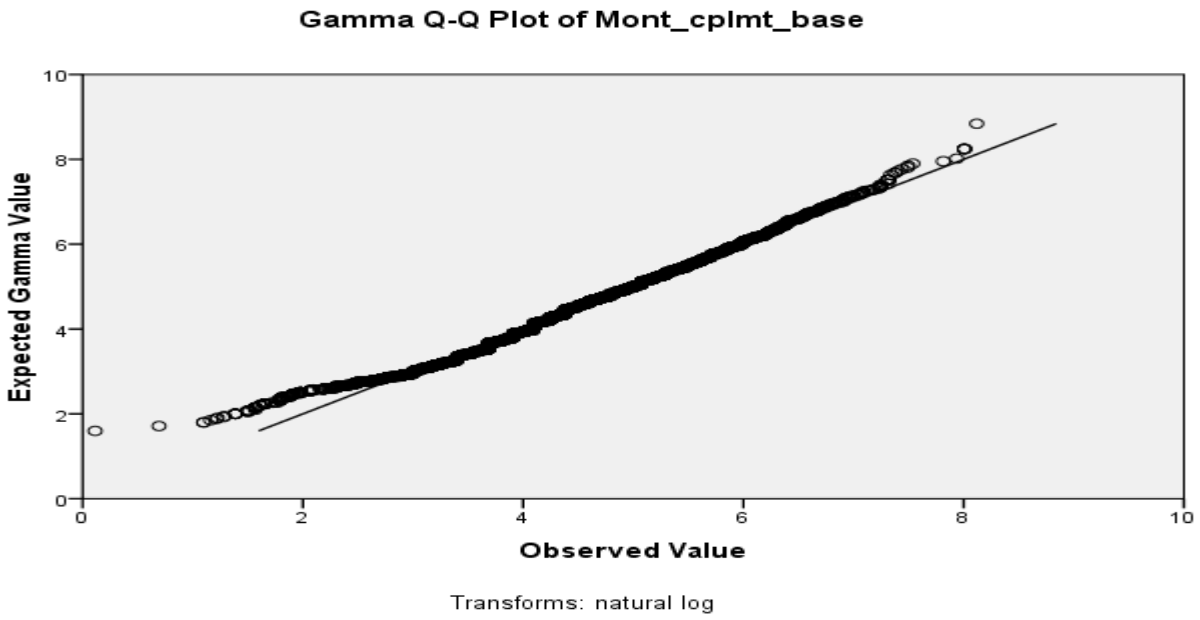

Figure 3.10: Ajustement Gamma de la transformation logarithmique des montants de sinistre.

Corrélations

		Taillecode	agecode	PLafondcode
Taillecode	Pearson Correlation	1,000	,592**	,604**
	Sig. (2-tailed)		,000	,000
	N	113,000	113	113
agecode	Pearson Correlation	,592**	1,000	,436**
	Sig. (2-tailed)	,000		,000
	N	113	113,000	113
PLafondcode	Pearson Correlation	,604**	,436**	1,000
	Sig. (2-tailed)	,000	,000	
	N	113	113	113,000

Tableau 3.33: Analyse de la corrélation de Pearson pour le poste Soins dentaires.

Annexe 11 : **Résultats de la première approche pour le poste Soins et Prothèses dentaires.**

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	1077,054	1	,000
Taillecode	134,040	4	,000

Tableau 3.34 : Analyse de type III de la variable Taillecode du poste Soins dentaires.

Goodness of Fit[b]

	Value	df	Value/df
Deviance	74,659	108	,691
Scaled Deviance	74,659	108	
Pearson Chi-Square	117,072	108	1,084
Scaled Pearson Chi-Square	117,072	108	
Log Likelihood[a]	-489,502		
Akaike's Information Criterion (AIC)	989,005		
Finite Sample Corrected AIC (AICC)	989,565		
Bayesian Information Criterion (BIC)	1002,642		
Consistent AIC (CAIC)	1007,642		

Dependent Variable: Fréquence

Model: (Intercept), Taillecode

Tableau 3.35 : indicateurs d'ajustement de la fréquence du poste Soins dentaires.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	6,261	,4476	5,384	7,139	195,658	1	,000	524,000	217,921	1259,979
[Taillecode=1]	-3,923	,4698	-4,843	-3,002	69,720	1	,000	,020	,008	,050
[Taillecode=2]	-3,102	,5082	-4,098	-2,106	37,257	1	,000	,045	,017	,122
[Taillecode=3]	-1,741	,5009	-2,723	-,760	12,084	1	,001	,175	,066	,468
[Taillecode=4]	-1,980	,5136	-2,987	-,974	14,872	1	,000	,138	,050	,378
[Taillecode=5]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									
(Negative binomial)	1									

Dependent Variable: Fréquence
Model: (Intercept), Taillecode

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Tableau 3.36 : Estimation des paramétres du modèle de la fréquence du poste Soins dentaires.

Estimated Distribution Parameters

		Mont_cplmt_base
Gamma Distribution	Shape	17,872
	Scale	3,284

Tableau 3.37: Estimation des paramètres dans la modélisation du cout des sinistres pour le poste Soins dentaires.

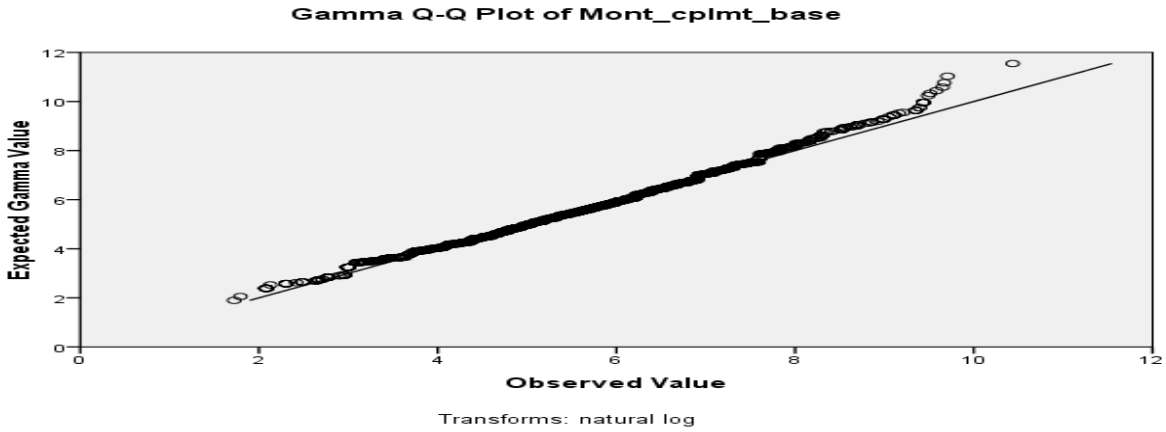

Figure 3.11: Ajustement Gamma de la transformation logarithmique des montants de sinistre

Annexe 12 : Résultats de la première approche pour le poste Optique.

Corrélations

		Taillecode	agecode	PLafondcode
Taillecode	Pearson Correlation	1,000	,566**	,608**
	Sig. (2-tailed)		,000	,000
	N	111,000	111	111
agecode	Pearson Correlation	,566**	1,000	,427**
	Sig. (2-tailed)	,000		,000
	N	111	111,000	111
PLafondcode	Pearson Correlation	,608**	,427**	1,000
	Sig. (2-tailed)	,000	,000	
	N	111	111	111,000

Tableau 3.38: Analyse de la corrélation de Pearson pour le poste Optique.

Goodness of Fit[b]

	Value	df	Value/df
Deviance	60,861	106	,574
Scaled Deviance	60,861	106	
Pearson Chi-Square	95,389	106	,900
Scaled Pearson Chi-Square	95,389	106	
Log Likelihood[a]	-453,473		
Akaike's Information Criterion (AIC)	916,946		
Finite Sample Corrected AIC (AICC)	917,517		
Bayesian Information Criterion (BIC)	930,494		
Consistent AIC (CAIC)	935,494		

Dependent Variable: frequence

Model: (Intercept), Taillecode

Tableau 3.39 : indicateurs d'ajustement de la fréquence du poste Optique.

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	906,694	1	,000
Taillecode	117,118	4	,000

Tableau 3.40 : Analyse de type III de la variable Taillecode du poste Optique.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	5,744	,4479	4,866	6,622	164,421	1	,000	312,200	129,764	751,123
[Taillecode_first=1]	-3,613	,4713	-4,536	-2,689	58,749	1	,000	,027	,011	,068
[Taillecode_first=2]	-2,930	,5094	-3,929	-1,932	33,084	1	,000	,053	,020	,145
[Taillecode_first=3]	-1,600	,5014	-2,583	-,617	10,178	1	,001	,202	,076	,540
[Taillecode_first=4]	-1,642	,5140	-2,649	-,635	10,206	1	,001	,194	,071	,530
[Taillecode_first=5]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									
(Negative binomial)	1									

Dependent Variable: frequence
Model: (Intercept), Taillecode_first

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Tableau 3.41. : Estimation des paramètres du modèle de la fréquence du poste Optique.

Estimated Distribution Parameters

		Mont_cplmt_base
Gamma Distribution	**Shape**	**29,151**
	Scale	**5,426**

Tableau 3.42: Estimation des paramètres dans la modélisation du cout des sinistres pour le poste Optique.

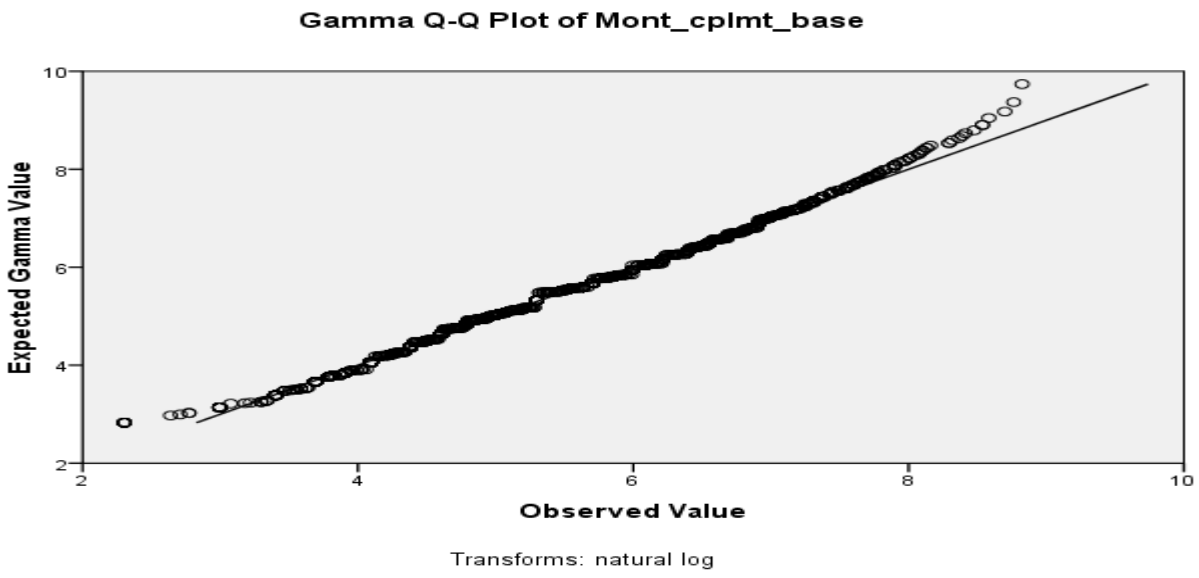

Figure 3.12: Ajustement Gamma de la transformation logarithmique des montants de sinistre.

Annexe 13 : Résultats de la première approche pour le poste Autres Soins.

Corrélations

		Taillecode	PLafondcode	agecode
Taillecode	Pearson Correlation	1,000	,554**	,622**
	Sig. (2-tailed)		,000	,000
	N	66,000	66	66
PLafondcode	Pearson Correlation	,554**	1,000	,473**
	Sig. (2-tailed)	,000		,000
	N	66	66,000	66
agecode	Pearson Correlation	,622**	,473**	1,000
	Sig. (2-tailed)	,000	,000	
	N	66	66	66,000

Tableau 3.43: Analyse de la corrélation de Pearson pour le poste Autres Prestations.

Goodness of Fit[b]

	Value	df	Value/df
Deviance	25,965	61	,426
Scaled Deviance	25,965	61	
Pearson Chi-Square	22,534	61	,369
Scaled Pearson Chi-Square	22,534	61	
Log Likelihood[a]	-177,920		
Akaike's Information Criterion (AIC)	365,839		
Finite Sample Corrected AIC (AICC)	366,839		
Bayesian Information Criterion (BIC)	376,787		
Consistent AIC (CAIC)	381,787		

Dependent Variable: frequence

Model: (Intercept), Taillecode

Tableau 3.44 : indicateurs d'ajustement de la fréquence du poste Autres.

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	134,071	1	,000
Taillecode	33,858	4	,000

Tableau 3.45 : Analyse de type III de la variable Taillecode du poste Autres Prestations.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	2,901	,4593	2,001	3,802	39,899	1	,000	18,200	7,397	44,777
[Taillecode=1]	-2,445	,5449	-3,513	-1,377	20,126	1	,000	,087	,030	,252
[Taillecode=2]	-1,515	,5915	-2,674	-,356	6,561	1	,010	,220	,069	,701
[Taillecode=3]	-,493	,5211	-1,515	,528	,897	1	,344	,611	,220	1,695
[Taillecode=4]	-1,266	,5391	-2,322	-,209	5,512	1	,019	,282	,098	,811
[Taillecode=5]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									
(Negative binomial)	1									

Dependent Variable: frequence
Model: (Intercept), Taillecode

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Tableau 3.45. : Estimation des paramètres du modèle de la fréquence du poste Autres.

Estimated Distribution Parameters

		Mont_cplmt_base
Gamma Distribution	Shape	15,261
	Scale	2,747

Tableau 3.46: Estimation des paramètres dans la modélisation du cout des sinistres pour le poste Autres.

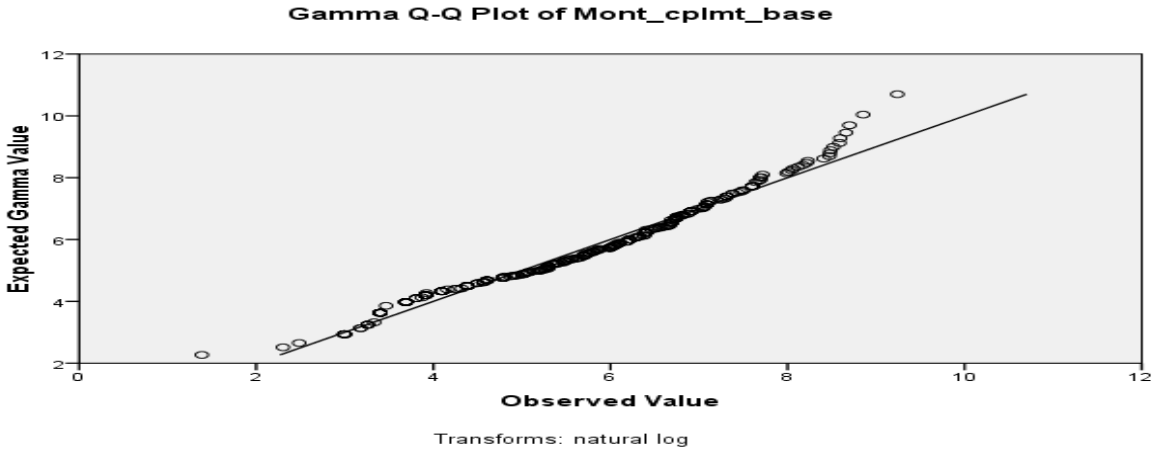

Figure 3.13: Ajustement Gamma de la transformation logarithmique des montants de sinistre.

Annexe 14 : Résultats de la deuxième approche de tarification pour le poste Ambulatoire.

[redacted]

Tableau 3.47: Sélection des variables du modèle Logit pour le poste Ambulatoire.

[redacted]

Tableau 3.48. : Estimation des paramètres de la régression logistique du poste ambulatoire.

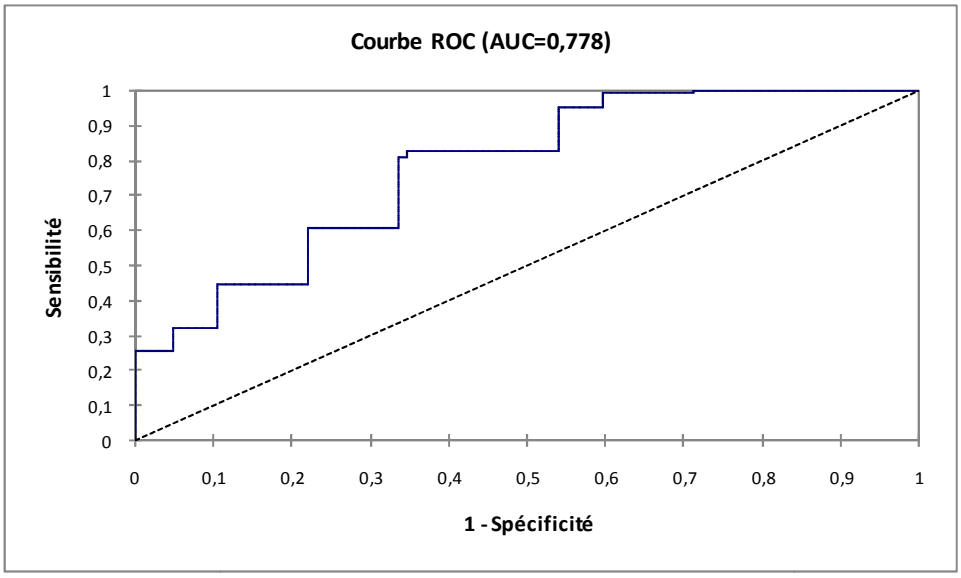

Figure 3.14: Courbe ROC du modèle Logit pour le poste Ambulatoire.

Paramètres estimés :

Paramètre	Valeur
μ	7,826
sigma	1,914

Tableau 3.48. : Estimation des paramètres de l'ajustement de la charge annuelle de sinistralité dans le poste ambulatoire.

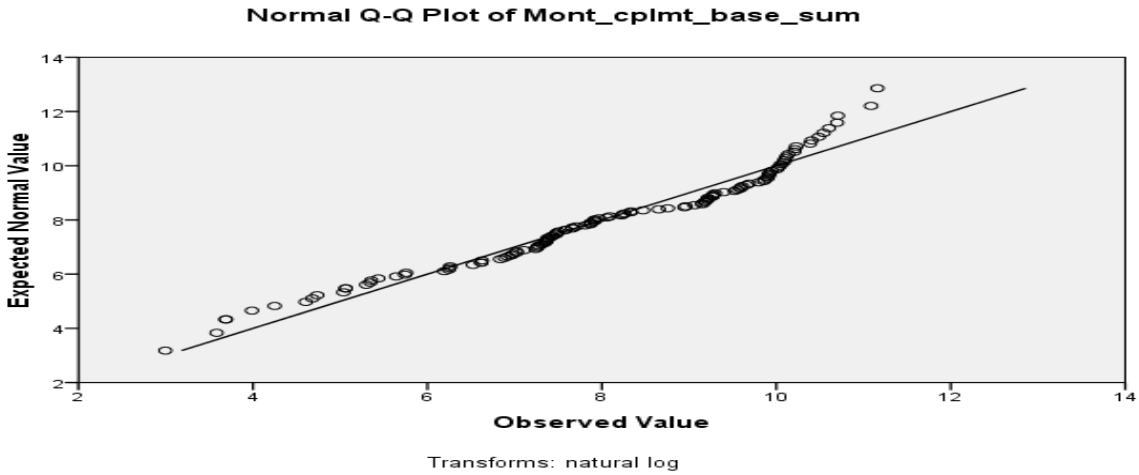

Figure 3.15: Ajustement Normal de la transformation logarithmique de la charge annuelle de sinistralité

Annexe 15 : Résultats de la deuxième approche de tarification pour le poste Hospitalisation.

[redacted]

Tableau 3.49: Sélection des variables du modèle Logit pour le poste Hospitalisation.

[redacted]

Tableau 3.50 : Estimation des paramètres de la régression logistique du poste Hospitalisation.

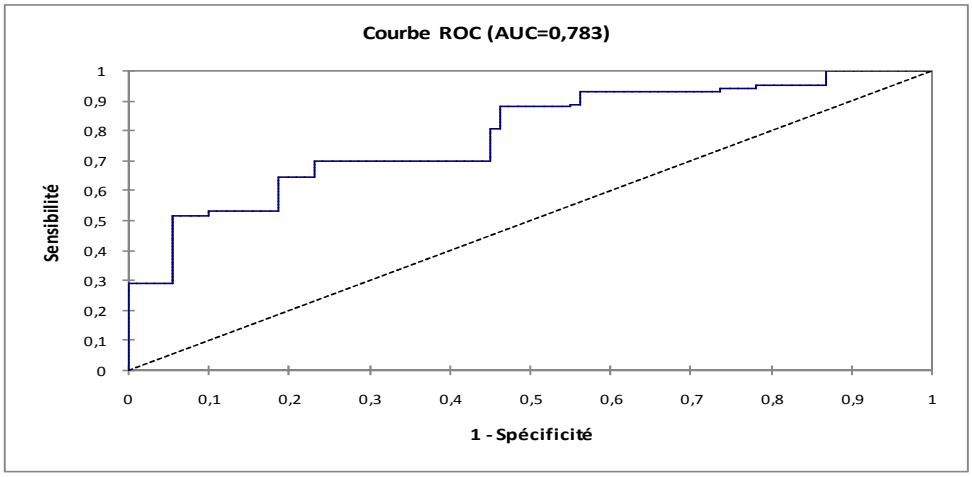

Figure 3.16: Courbe ROC du modèle Logit pour le poste Hospitalisation.

Paramètres estimés :	
Paramètre	**Valeur**
µ	**8,144**
sigma	**1,911**

Tableau 3.51: Estimation des paramètres de l'ajustement de la charge annuelle de sinistralité dans le poste Hospitalisation.

Test de Kolmogorov-Smirnov	
D	**0,073**
p-value	**0,657**
alpha	**0,05**

Tableau 3.52: Test d'adéquation de la loi Normale à la charge annuelle de sinistralité dans le poste Hospitalisation.

Interprétation du test :

H0 *: l'échantillon suit la loi normale.*

H1 *: l'échantillon ne suit pas la loi normale.*

Etant donné que la p-value du test est supérieure à alpha=0.005, on peut valider H0.

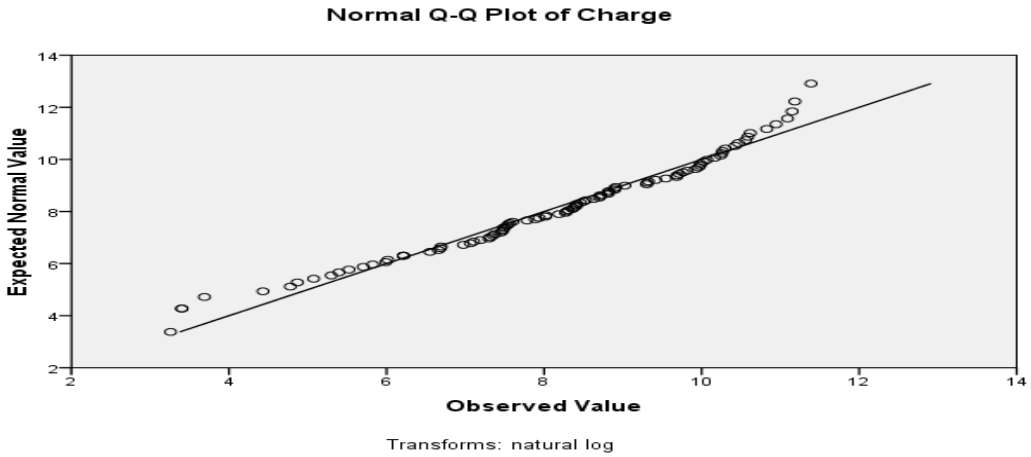

Figure 3.17: Ajustement Normal de la transformation logarithmique de la charge annuelle de sinistralité.

Annexe 16 : Résultats de la deuxième approche de tarification pour le poste Analyses et Radios.

[redacted]

Tableau 3.53: Sélection des variables du modèle Logit pour le poste Radios.

[redacted]

Tableau 3.54 : Estimation des paramètres de la régression logistique du poste Radios.

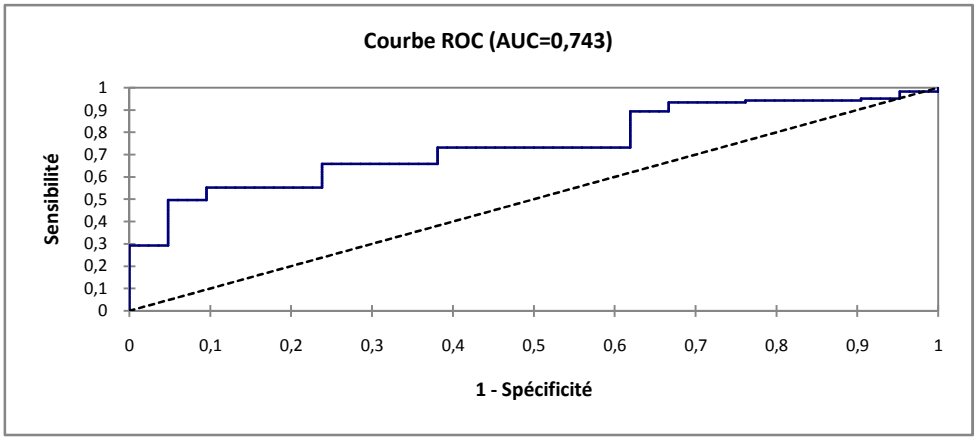

Figure 3.18: Courbe ROC du modèle Logit pour le poste Radios.

Paramètres estimés :	
Paramètre	**Valeur**
μ	**8,029**
sigma	**1,880**

Tableau 3.55: Estimation des paramètres de l'ajustement de la charge annuelle de sinistralité dans le poste Radios.

Test de Kolmogorov-Smirnov :	
D	**0,070**
p-value	**0,561**
alpha	**0,05**

Tableau 3.56: Test d'adéquation de la loi Normale à la charge annuelle de sinistralité dans le poste Radios.

Interprétation du test :

H0 *: l'échantillon suit la loi normale.*

H1 *: l'échantillon ne suit pas la loi normale.*

Etant donné que la p-value du test est supérieure à alpha=0.005, on peut valider H0.

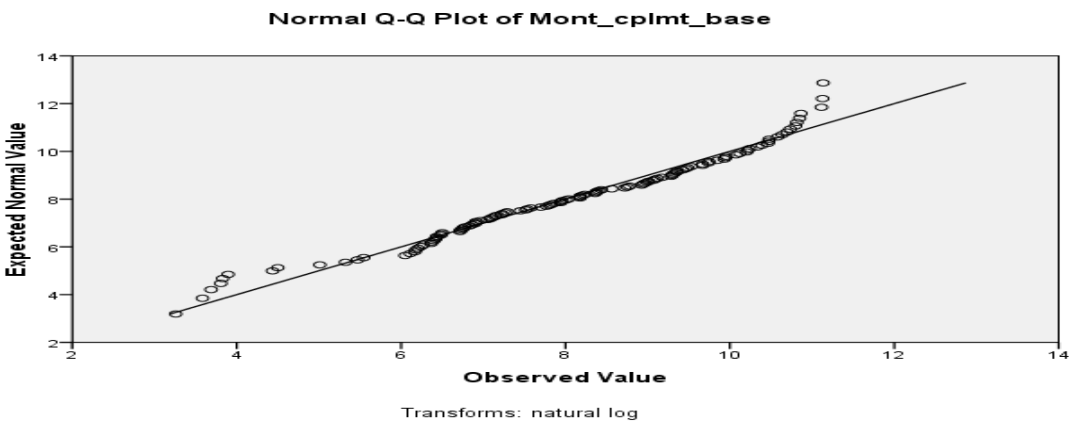

Figure 3.19: Ajustement Normal de la transformation logarithmique de la charge annuelle de sinistralité.

Annexe 17 : Résultats de la deuxième approche de tarification pour le poste Soins dentaires.

[redacted]

Tableau 3.57: Sélection des variables du modèle Logit pour le poste Soins dentaires.

[redacted]

Tableau 3.58 : Estimation des paramètres de la régression logistique du poste Soins dentaires.

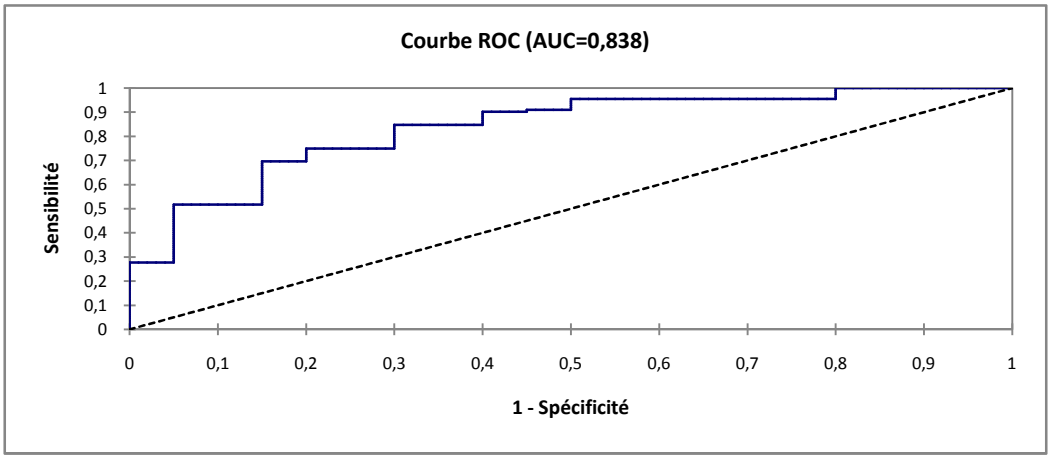

Figure 3.20: Courbe ROC du modèle Logit pour le poste Soins dentaires.

Paramètres estimés :	
Paramètre	**Valeur**
μ	**9,184**
sigma	**1,653**

Tableau 3.59: Estimation des paramètres de l'ajustement de la charge annuelle de sinistralité dans le poste Soins dentaires :

Test de Kolmogorov Smirnov	
D	**0,065**
p-value	**0,709**
alpha	**0,05**

Tableau 3.60: Test d'adéquation de la loi Normale à la charge annuelle de sinistralité dans le poste Radios.

Interprétation du test :

H0 *: l'échantillon suit la loi normale.*

H1 *: l'échantillon ne suit pas la loi normale.*

Etant donné que la p-value du test est supérieure à alpha=0.005, on peut valider H0.

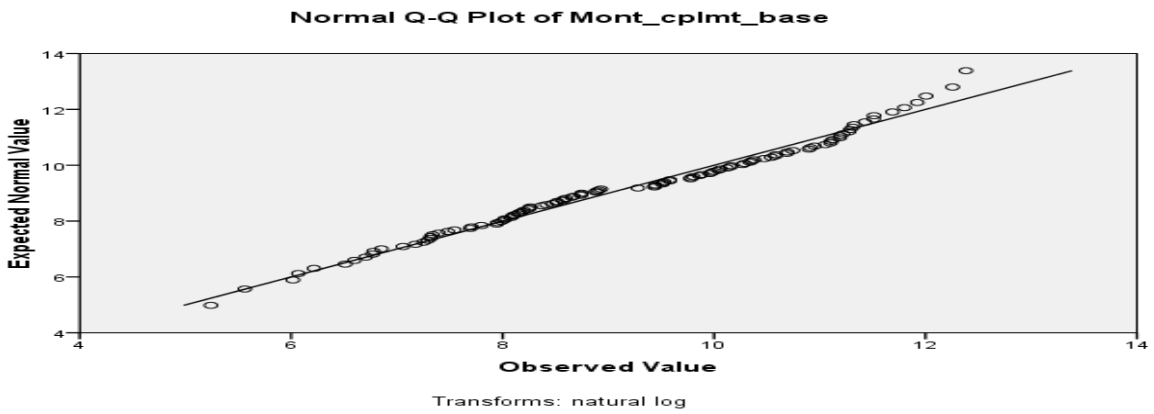

Figure 3.21: Ajustement Normal de la transformation logarithmique de la charge annuelle de sinistralité.

Annexe 18 : Résultats de la deuxième approche de tarification pour le poste Optique.

[redacted]

Tableau 3.61: Sélection des variables du modèle Logit pour le poste Optique.

[redacted]

Tableau 3.62 : Estimation des paramètres de la régression logistique du poste Optique.

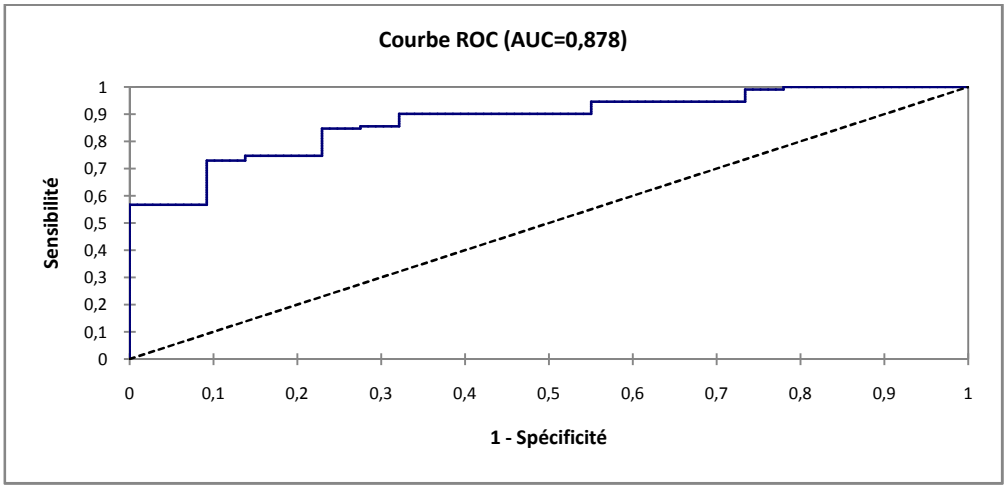

Figure 3.22: Courbe ROC du modèle Logit pour le poste Optique.

Paramètres estimés :	
Paramètre	**Valeur**
µ	**8,525**
sigma	**1,682**

Tableau 3.63: Estimation des paramètres de l'ajustement de la charge annuelle de sinistralité dans le poste Optique.

Test de Kolmogorov-Smirnov :	
D	**0,072**
p-value	**0,594**
alpha	**0,05**

Tableau 3.64: Test d'adéquation de la loi Normale à la charge annuelle de sinistralité dans le poste Optique.

Interprétation du test :

H0 *: l'échantillon suit la loi normale.*

H1 *:l'échantillon ne suit pas la loi normale.*

Etant donné que la p-value du test est supérieure à alpha=0.005, on peut valider H0.

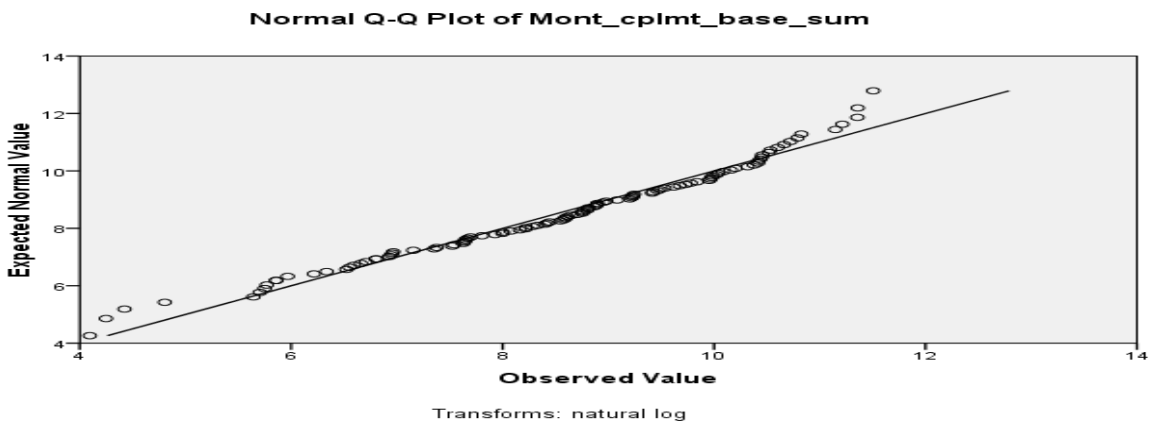

Figure 3.23: Ajustement Normal de la transformation logarithmique de la charge annuelle de sinistralité.

Annexe 19 : Résultats de la deuxième approche de tarification pour le poste Autres prestations.

[redacted]

Tableau 3.64: Sélection des variables du modèle Logit pour le poste Autres.

[redacted]

Tableau 3.65 : Estimation des paramètres de la régression logistique du poste Autres.

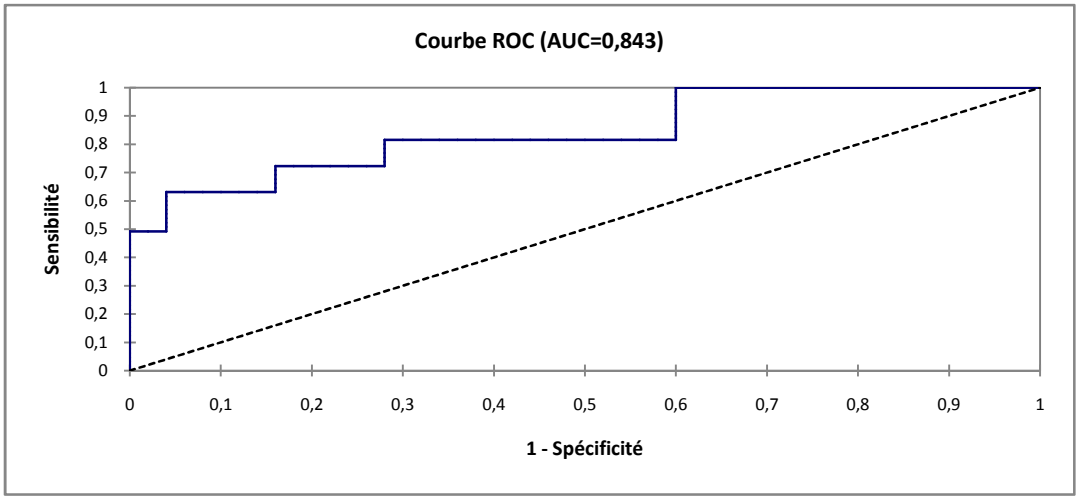

Figure 3.24: Courbe ROC du modèle Logit pour le poste Autres.

Paramètres estimés :	
Paramètre	**Valeur**
µ	**7,082**
sigma	**1,893**

Tableau 3.66: Estimation des paramètres de l'ajustement de la charge annuelle de sinistralité dans le poste Autres.

Test de Kolmogorov-Smirnov :	
D	**0,114**
p-value	**0,338**
alpha	**0,05**

Tableau 3.67: Test d'adéquation de la loi Normale à la charge annuelle de sinistralité dans le poste Autres.

Interprétation du test :

H0 *: l'échantillon suit la loi normale.*

H1 *:l'échantillon ne suit pas la loi normale.*

Etant donné que la p-value du test est supérieure à alpha=0.005, on peut valider H0.

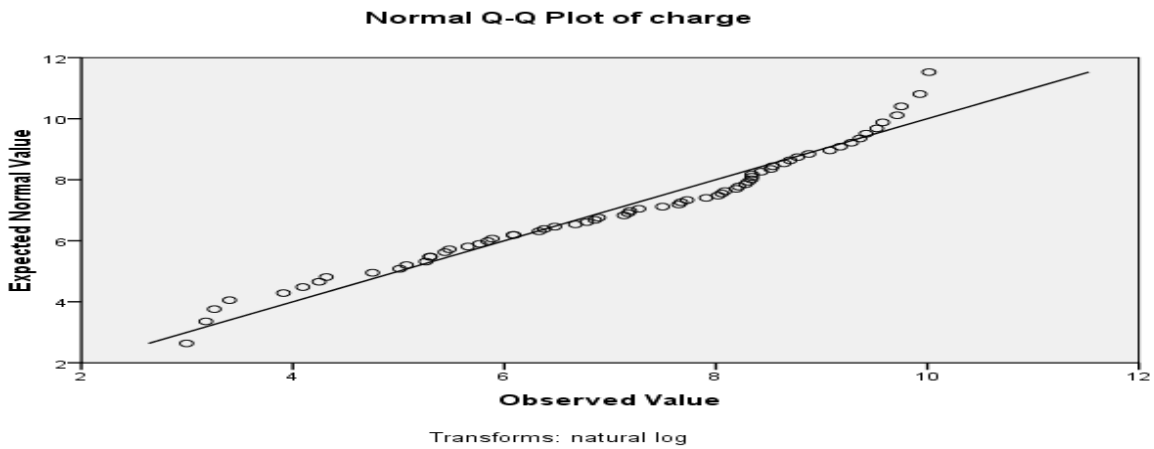

Figure 3.25: Ajustement Normal de la transformation logarithmique de la charge annuelle de sinistralité.