

INSTITUT NATIONAL
DE STATISTIQUE ET D'ECONOMIE APPLIQUEE

INSEA



Projet de Fin d'Etudes

Evaluation des engagements d'AXA dans le portefeuille de la CIMR

Préparé par : *M. EL GRINY Bouchaib*
M. MOAHAMED AHID Sidi Abdellahi

Sous la direction de : *M. MARRI Fouad (INSEA)*
M. MOUATASSIM Youness (AXA Assurance Maroc)

*Soutenu publiquement comme exigence partielle en vue de l'obtention du
Diplôme d'Ingénieur d'Etat*

Option : Actuariat-Finance

Devant le jury composé de :

- *M. MARRI Fouad (INSEA)*
- *M. EL QALLI Yassine (INSEA)*
- *M. MOUATASSIM Youness (AXA Assurance Maroc)*

Juin 2014

Dédicaces

A la mémoire de mon père.

À ma mère, pour son dévouement, ses sacrifices et son amour,

À mes cousines et ma tante Fatima, pour leurs encouragements et leur soutien,

Au corps professoral de l'INSEA, pour ses efforts et sa rigueur,

À mon binôme, pour son esprit d'équipe, sa motivation, sa persévérance et son dynamisme,

À mes camarades de l'INSEA, pour leur solidarité et leur appui tout au long de ces trois années de formation, particulièrement, Rabab, l'barhmi, mido, mouki, toubi et l'mahdani, pour leur esprit d'initiative et leur sens de collaboration.

Au personnel de la Direction du Marché des Particuliers : Adil, Hamid, Fatima Ezzahra, Najoua et Souad.

EL GRINY Bouchaib

A mes très chers et estimables parents qui n'ont cessé de me guider vers le droit chemin ainsi pour leurs sacrifices et leur amour permanent.

À la mémoire de ma grand-mère,

À mes chers frères,

À ma chère petite,

A mes oncles, tantes et cousins,

A mon binôme très studieux, enthousiasmé et dynamique.

Au personnel de la Direction du Marché des Particuliers : Adil, Hamid, Fatima Ezzahra, Najoua et Souad.

A mes chères amis compatriotes, marocains et ceux de l'Afrique subsaharienne pour leur énorme apport et soutien.

Au Royaume du Maroc.

SIDI Abdellatif

Remerciements

Il va sans dire que nul ne peut mener à bien l'élaboration de son Projet de Fin d'Etudes sans l'appui des encadrants. Ces personnes expérimentées, dotées d'un esprit d'analyse et d'un sens de responsabilité, qui n'épargnent aucun effort pour que nous puissions appréhender et répondre à la problématique du PFE et envers lesquels ne pouvant qu'éprouver une grande estime et une profonde reconnaissance. Nous tenons donc à adresser nos vifs remerciements à notre encadrant interne, Monsieur MARRI Fouad, pour sa disponibilité, son soutien et ses conseils avisés. Tout comme nous tenons à remercier Monsieur MOUATASSIM Youness pour la pertinence de ses remarques et sa bienveillance tout au long de notre période de stage au sein d'AXA Assurance Maroc.

Nous tenons aussi à exprimer notre profonde gratitude envers le corps professoral de l'INSEA, particulièrement Monsieur TOUHAMI Abdelkhalék et Monsieur CHAOUBI Abdelaziz pour leurs conseils et leurs remarques à forte valeur ajoutée.

Enfin, nous remercions Monsieur DBICH Abderrahim pour son offre de stage, sa rigueur et son attention, ainsi que tout le personnel d'AXA Assurance Maroc, particulièrement celui du département « Assurances Dommages » qui nous a garanti, par son aisance relationnelle et son esprit d'initiative, une atmosphère agréable et chaleureuse.

Que toutes les personnes qui ont contribué, de près ou de loin, à l'élaboration de ce PFE trouvent dans ce remerciement l'expression de notre profonde reconnaissance.

Résumé

Dans le cadre du contrat régime commun conclu entre les compagnies d'assurances représentées légalement par la Fédération Marocaine des Sociétés d'Assurances et de Réassurance et la Caisse Interprofessionnelle Marocaine de Retraite (CIMR). Les affiliés de la CIMR possèdent deux comptes retraite, un premier géré par la CIMR en répartition et un autre géré par la compagnie d'assurance en charge par capitalisation. AXA Assurance Maroc gère une part assez importante de ce portefeuille.

Dans ce contexte, AXA se doit toujours d'être alertée quant aux comportements futurs de ces assurés lors de la liquidation de leurs retraites. Des comportements qui ne sont pas toujours les mêmes et qui méritent un traitement particulier.

Dans ce mémoire, nous avons essayé d'approcher le comportement de chacun des contrats en cours afin de pouvoir estimer la réserve globale qui s'en déduit. A cet effet, nous avons modélisé le portefeuille à l'aide de la régression logistique, selon l'option de sortie en capital et l'option de sortie en rachat et aussi à l'aide de l'arbre de décision, pour les options contre-assurance et rente viagère.

Par la suite, nous avons pu faire les prédictions en se basant sur les modèles estimés. Nous avons pu aussi identifier le comportement probable des assurés ayant un contrat en cours. Enfin, nous avons procédé aux calculs des réserves mathématiques pour chacun des types de sortie.

Mots clés

Régime Commun par capitalisation, Réserves mathématiques, Rente viagère, Rente réversible, Prorogation, anticipation, Engagement, Loi de sortie, logistique, arbre de décision.

Table des matières

Dédicaces	2
Remerciements	3
Résumé	4
Mots clés	4
Liste des acronymes	7
Liste des graphiques	8
Liste des tableaux	8
Introduction générale	9
Chapitre 1 : Système de retraite au Maroc	11
Introduction	11
1. Généralités sur le système de retraite	11
1.1 Définition de la sécurité sociale	11
1.2 Définition du concept « retraite »	12
1.3 Concept du régime de retraite	12
2. Système de retraite marocain	13
2.1 Montage institutionnel des régimes de retraite au Maroc	13
2.2 Régime de retraite de la Caisse Interprofessionnelle Marocaine de Retraite	14
2.3 L'assurance retraite selon la vision d'AXA Assurance Maroc	19
Conclusion	20
Chapitre 2 : Certification du système interne d'AXA	21
Introduction	21
1. Cadre légal Pour les paramètres de calcul	21
1.1 Tarifs	21
1.2 Provisionnement	22
2. Méthodologie	22
2.1 Calcul des cotisations et des rentes	22
2.2 Calcul des coefficients d'anticipation et de prorogation	24
2.3 Calcul des réserves mathématiques	25
Conclusion	27
Chapitre 3 : Modélisation de la loi de sortie	28
Introduction	28
1. Présentation et Préparation de la Base de données	29

2. Mise en pratique des techniques statistiques	30
2.1 Description du portefeuille	31
2.2 Cadre théorique du modèle logistique	34
2.3 Arbres de décision	41
2.4 Résultats	44
3. Evaluation des Engagements AXA	61
3.1 Hypothèses	61
3.2 Résultats	63
Conclusion	64
Conclusion générale et perspectives	65
Références bibliographiques	67
Annexes	68

Liste des acronymes

<i>CIMR</i>	:	Caisse Interprofessionnelle Marocaine de Retraite
<i>CAT</i>	:	Coefficient à terme
<i>CA</i>	:	Cotisation Annuelle
<i>R_a</i>	:	Rente acquise
<i>R_r</i>	:	Rente Réversible
<i>TC</i>	:	Taux de cotisation
<i>S</i>	:	Salaire annuel
<i>V_x</i>	:	Réserves Mathématiques
<i>CHAID</i>	:	Chi Squared Automatic Interaction Detector
<i>CAP</i>	:	Sortie en Capital
<i>CAS</i>	:	Sortie par contre-assurance
<i>RVR</i>	:	Sortie en rente viagère
<i>RAC</i>	:	Sortie par rachat
<i>AIC</i>	:	Akaic Information Criterion
<i>BIC</i>	:	Bayesian Information Criterion
<i>OR</i>	:	Odds ratio
<i>CFG</i>	:	Chargements pour frais de gestion
<i>CFA</i>	:	Chargement pour frais d'acquisition

Liste des graphiques

Graphique 1 : Répartition de la sortie en fonction des exercices.....	32
Graphique 2 : Répartition des fréquences de la loi de sortie	32
Graphique 3 : Fréquence de sortie en fonction de l'âge et du type de sortie.....	33
Graphique 4 : Répartition de l'historique en fonction de l'ancienneté	33
Graphique 5 : Répartition de l'âge à l'embauche en fonction des tranches du cumul de cotisation.....	46
Graphique 6 : Répartition d'ancienneté en fonction de tranche cotisation.....	46
Graphique 7 : Répartition du logit(p) en fonction de l'option capital	51
Graphique 8 : La courbe ROC du modèle de l'option capital	51
Graphique 9 : Répartition du logit(p) en fonction de l'option rachat	54
Graphique 10 : La courbe ROC pour l'option rachat	55
Graphique 11 : La courbe de gain du modèle développé par l'arbre de décision.	60
Graphique 12 : La courbe Index de l'arbre de décision.	60

Liste des tableaux

Tableau 1 : techniques statistiques de prédiction	31
Tableau 2 : répartition des sorties en fonction du sexe	34
Tableau 3 : la matrice de corrélation de Pearson.....	45
Tableau 4 : Coefficient de corrélation pour chaque couple des variables	48
Tableau 5 : Résultats du model sous SAS.....	48
Tableau 6 : test de signification globale	49
Tableau 7 : Estimation des coefficients du modèle de capital.....	49
Tableau 8 : Statistiques d'ajustement du modèle de capital	50
Tableau 9 : Table de confusion pour le modèle du capital	50
Tableau 10 : statistiques d'ajustement du modèle de rachat	52
Tableau 11 : Tests d'ajustements du modèle.....	53
Tableau 12 : Estimation des coefficients du modèle de capital.....	53
Tableau 13 : Tableau croisé dressant les valeurs observées contre les valeurs prédites	54
Tableau 14 : Croisement des valeurs observées de l'historique des options rachat et capital	55
Tableau 15 : Croisement des valeurs prédites sur l'historique des options rachat et capital ..	56
Tableau 16 : Croisement des valeurs prédites de l'option de sortie des contrats en cours	57
Tableau 17 : Test d'ajustement de l'arbre des échantillons	59
Tableau 18 : Matrice de confusion pour les échantillons d'apprentissage et du test	59
Tableau 19 : Résultats de la prévision via le modèle pour la rente viagère et contre-assurance	61
Tableau 20 : Formules utilisées pour le calcul des réserves pour les âges inférieures à 60 ans.	62
Tableau 21 : Formules utilisées pour le calcul des réserves pour les âges inférieures à 60 ans.	63
Tableau 22 : Répartition des réserves selon les types de sorties.	63

Introduction générale

Contexte général et problématique :

La prise en compte des risques de l'existence (risques sociaux), différemment par les Etats, était un tournant radical dans l'histoire de l'Homme; la question de leurs couvertures a mené à la création du système de protection sociale susceptible d'être assuré par un ensemble de techniques quelle que soit la nature des orientations (principe de solidarité et principe de responsabilité).

En effet, la protection sociale désigne toutes politiques et actions qui permettent aux pauvres et aux personnes vulnérables de mieux faire face aux situations à risque et aux chocs et qui leur donnent davantage de moyens pour sortir du dénuement. Elle recouvre l'assurance sociale, les transferts sociaux et les normes minimales du travail.

De plus, l'existence d'une économie forte et équitable, à laquelle les femmes et les hommes sont directement associés tant parce qu'ils contribuent que parce qu'ils en profitent, est une condition essentielle de la réduction de la pauvreté.

Aussi, les défis permanents tels l'évolution de la population, l'instabilité des prix et l'insécurité alimentaire, mettant en évidence la vulnérabilité des individus et familles pauvres, exigent aux pouvoirs publics de prendre des mesures plus efficaces pour remédier aux difficultés existantes.

Ainsi, la protection sociale, notamment par le biais de la sécurité sociale ou assurance sociale (retraite et autres), est considérée comme l'un des moyens déterminants pour assurer l'instauration d'une croissance pro-pauvre et la réalisation des objectifs du millénaire de développement. C'est dans ce souci de croissance qu'apparaît le rôle primordial pour la survie et le développement que joue la retraite ou assurance de groupe.

En fait, L'équilibre budgétaire de certains organismes de la protection sociale, pose de façon aiguë le problème de son devenir et ce, en remettant en cause les fondements de base de la pérennité du système.

C'est dans ce sens que toute compagnie d'assurance ou organisme financier ayant pour vocation de gérer les contributions (cotisations) de ses adhérents (assurés), auxquels elle appartient la société d'assurance AXA, est tenue par la législation de constituer des

réserves, appelées aussi provisions, afin de faire face à tout moment à ses engagements vis-à-vis de ses assurés. C'est dans ce cadre que la réévaluation des engagements d'AXA Assurance Maroc dans le portefeuille de la Caisse Interprofessionnelle Marocaine de Retraite (CIMR) fera l'objet de notre étude.

Démarche :

Pour mener cette question d'évaluation des engagements d'AXA dans le portefeuille de la CIMR, nous avons scindé ce travail en trois chapitres comme suit :

Dans un premier chapitre, il sera présenté une introduction du montage institutionnel du système de retraite marocain dans sa globalité et notamment celui de la CIMR, suivie d'une présentation chiffrée des principaux indicateurs réalisés par AXA Assurance Maroc. Dans un deuxième chapitre, il sera abordé la question de la certification du système interne d'AXA. Dans ce chapitre nous présenterons la phase de reproduction du système par le biais des bases techniques et par le calcul des tarifs, des provisions, des coefficients d'anticipation et de prorogation ainsi que des réserves mathématiques. La réalisation de ces calculs n'était pas possible sans la préparation et le traitement préalable d'une base de données adéquate. Cette dernière a été mise en place sur la base d'un travail d'élaboration statistique pointu que nous avons mené depuis notre premier jour de stage jusqu'à l'achèvement de ce rapport.

Les définitions des concepts utilisés et les calculs réalisés dans le cadre de ces deux premiers chapitres ont servi à une modélisation de la loi de sortie des assurés du portefeuille et ce, en utilisant la méthode de segmentation selon une approche d'analyse des données, à savoir la régression logistique pour mieux évaluer les engagements. En se basant sur les hypothèses traduites par les résultats de la modélisation de la loi de sortie, il sera procédé par la projection des engagements (réserves).

Enfin, des recommandations seront émises, en guise de conclusion, avec des perspectives d'évolution qui peuvent faire l'objet d'autres travaux de recherche.

Pour alléger le document et faciliter sa lecture, certains détails des calculs et des résultats des techniques décrites dans ce travail sont présentés en annexe.

Chapitre 1 : Système de retraite au Maroc

INTRODUCTION

Adopter des politiques de développement visant à assurer un minimum de ressources à tous les citoyens exerçant ou ayant exercé une activité dans un cadre de sérénité, ainsi que l'accès aux besoins essentiels de la vie (éducation, nourriture ...), est un véritable défi auquel un système de retraite doit relever. Il fait également partie de la politique sociale, celle-ci étant intégrée dans la politique publique.

L'objectif de ce chapitre est de donner un aperçu global sur le système de retraite au Maroc et celui de la CIMR en particulier.

1. GENERALITES SUR LE SYSTEME DE RETRAITE

1.1 DEFINITION DE LA SECURITE SOCIALE

La sécurité sociale est une institution groupant divers organismes étatiques ou de droits privés chargés d'une mission de service public, auxquels la loi a conféré le soin de gérer des régimes obligatoires couvrant l'ensemble des risques sociaux des travailleurs salariés ou non-salariés et des professions libérales.

Elle a pour objectifs l'assurance, la réduction des inégalités et la lutte contre la pauvreté. Historiquement, il existe deux modèles de base de la sécurité sociale à savoir :

- Le modèle de Bismarck sur les assurances sociales en 1881 (Allemagne).
- Le modèle de Beveridge en 1942 (Angleterre).

La sécurité sociale : Un droit de l'Homme :

L'insertion de la sécurité sociale dans la déclaration universelle des Droits de l'Homme adoptée le 10 décembre 1948 par l'Assemblée générale des Nations Unies : « *Toute personne, en tant que membre de la société, a droit à la sécurité sociale; ...* » (Article 22);

« *Toute personne a droit à un niveau de vie suffisant pour assurer sa santé, son bien-être et ceux de sa famille (...); elle a droit à la sécurité en cas de chômage, de maladie, d'invalidité, de veuvage, de vieillesse*

ou dans les autres cas de perte de ses moyens de subsistance par suite de circonstances indépendantes de sa volonté» (Article 25).

1.2 DEFINITION DU CONCEPT « RETRAITE »

Le concept « retraite » est polysémique. Dans notre cas, il est défini comme étant la situation sociale et financière d'un individu qui, ayant atteint l'âge minimum requis, cesse son activité professionnelle, en général de manière définitive et perçoit régulièrement un revenu sous forme de pension versée par l'organisme chargé de la retraite ou de ressources provenant de ses cotisations passées ou de son patrimoine.

L'âge minimum de départ en retraite est soumis, selon les pays, à diverses conditions d'âge, de participation financière légale, d'état de santé, de maternité, d'handicap, d'invalidité....etc.

1.3 CONCEPT DU REGIME DE RETRAITE

La notion de régime de retraite, en tant que mécanisme systématique et organisé sur une base légale ou réglementaire d'octroi de prestations aux personnes âgées, est relativement récente ; il s'agit essentiellement de créations du XXe siècle.

Autrement dit, un régime de retraite est un ensemble de droits et obligations réciproques des employés (et leurs « ayants droits »), et des employeurs et d'un organisme de sécurité sociale.

On peut distinguer principalement les régimes de sécurité sociale mis en place au niveau de l'ensemble d'un pays, des régimes de retraite professionnels s'adressant à tous les membres d'une profession ou encore de plans complémentaires créés au sein des entreprises.

En fait, la structure des régimes des retraites repose sur trois composantes, complémentaires et de nature différente :

- Un premier pilier : la SECURITE SOCIALE organisée au niveau général d'un pays et permettant d'octroyer un premier niveau de base en matière de pension.
- un second pilier : les REGIMES PROFESSIONNELS organisés au sein d'une entreprise ou d'un secteur d'activité, octroyant à chacun des affiliés de ces régimes un complément à la sécurité sociale.

- un troisième pilier : l'ÉPARGNE INDIVIDUELLE organisée au libre choix de chacun.

2. SYSTEME DE RETRAITE MAROCAIN

Le système de retraite marocain couvre les salariés du secteur public et ceux du secteur privé. Il leur offre, également une assurance contre les divers risques tels ceux de santé, maternité, vieillesse, survie, décès et les sert des prestations familiales.

On peut distinguer essentiellement deux types de régimes au Maroc : les régimes de base et les régimes complémentaires.

Régimes de base

Un régime de base est un régime dont la vocation est de garantir une prestation minimale qui peut être forfaitaire ou variable selon le salaire de référence. Il est généralement conçu comme un régime auquel l'adhésion est obligatoire. En effet, l'affiliation cotisante à une caisse de retraite de base est obligatoire afin de bénéficier d'une prestation appelée pension de retraite. (Voir Annexes)

Régimes complémentaires

Comme son nom l'indique, un régime complémentaire est un régime dans lequel les prestations perçues représentent un complément à la retraite et pas un substitut à celle-ci.

Il a pour vocation de compléter et de poursuivre la prestation du régime de base avec comme objectif de rapprocher le niveau de la pension calculée de celui du dernier salaire perçu ou d'une fraction significative de ce dernier. (Voir Annexes)

2.1 MONTAGE INSTITUTIONNEL DES REGIMES DE RETRAITE AU MAROC

Il existe différents organismes de sécurité sociale au Maroc, il s'agit principalement de :

- Caisse Nationale de Sécurité Sociale (CNSS).
- Caisse Marocaine de Retraites (CMR).
- Régime Collectif d'Allocation de Retraites (RCAR).
- Caisse Interprofessionnelle Marocaine de Retraite (CIMR).
- Caisses internes telles celles de OCP, ONE, ODEP, ONCF (introduites à la RCAR) et d'autres "petits" régimes tels ceux de Bank Al Maghreb, des membres de la Chambre des représentants et de la Chambre des conseillers et le régime privé des compagnies d'assurance.

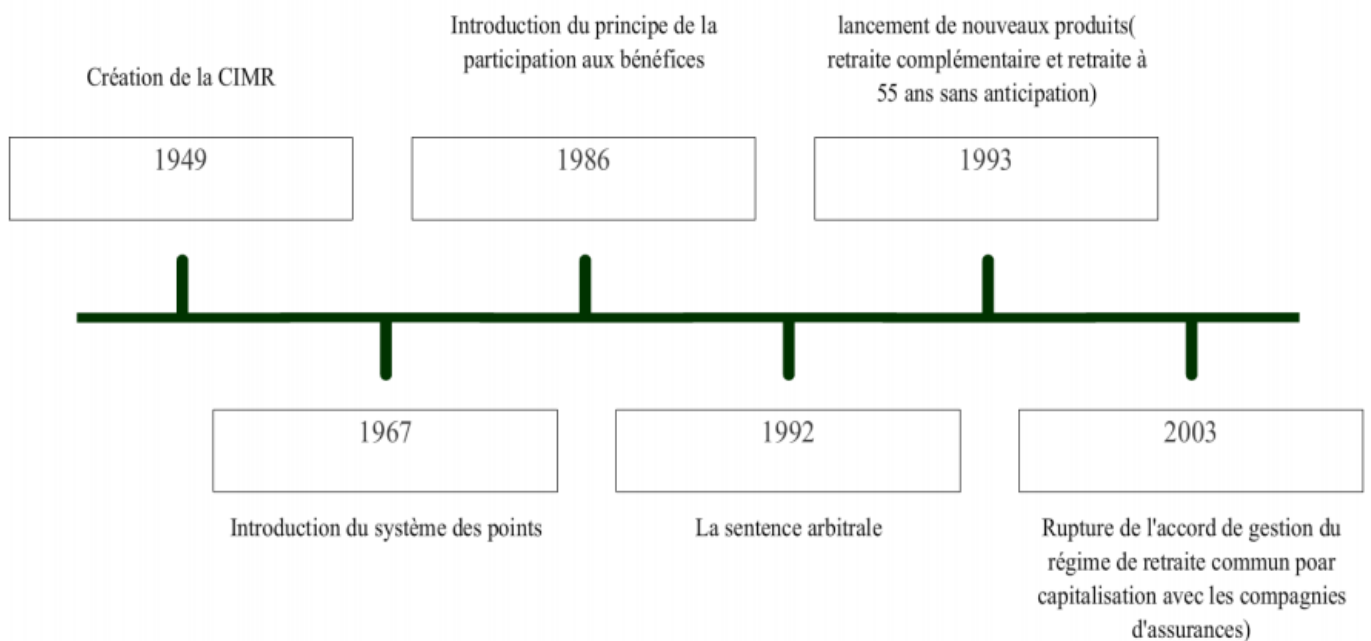
La Caisse Interprofessionnelle Marocaine de Retraite est d'une structure assez particulière. Elle présente des produits de retraite de base ainsi que des produits de retraite complémentaire, sachant qu'elle n'a pas un caractère obligatoire. En effet l'affiliation à cette caisse est purement facultative. Nous présentons ci-après, un focus sur cet organisme : son histoire et son mode de gestion.

2.2 REGIME DE RETRAITE DE LA CAISSE INTERPROFESSIONNELLE MAROCAINE DE RETRAITE

2.2.1 Aperçu sur la Caisse Interprofessionnelle Marocaine de Retraite

L'évolution de la Caisse Interprofessionnelle Marocaine de Retraite se présente comme suit :

Figure 1 : Faits marquants dans l'histoire de la Caisse Interprofessionnelle Marocaine de Retraite



- Création en 1949 sous forme d'association à but non lucratif, c'est le second organisme de prévoyance créé au Maroc après la Caisse Marocaine des Retraites et destiné au secteur privé et aux entreprises publiques et semi publiques.
- Au sein de la Caisse Interprofessionnelle Marocaine de Retraite (CIMR) coexistent deux régimes : le premier de retraite normale ou de base et le second de retraite complémentaire.
- La CIMR reçoit des primes qui se composent de deux parts, une part patronale qui la gère elle-même en répartition et une autre salariale confiée aux compagnies d'assurance pour la gérer en capitalisation dans le cadre d'un contrat commun

(régime commun) qui la lie avec la fédération Marocaine des sociétés d'assurances et de réassurances (FMSAR).

- La CIMR a introduit le système des points en 1968, pour lequel est maintenu jusqu'à aujourd'hui.
- La CIMR adopte un régime de cotisations définies, les prestations se calculent sur la base de l'ensemble des points acquis durant toute la période de service pour l'âge normale de retraite.
- Le taux minimum de contribution dans le cadre du régime de retraite normal est de 3% du traitement de base qui est déterminé au début de chaque année par la CIMR en fonction de la moyenne des salaires soumis à contribution. Toutefois les adhérents peuvent cotiser plus dans la limite de 6% du traitement de base.
- En 1986, la CIMR a introduit le système de participation aux bénéfices concernant le régime commun.
- En 1992, la sentence arbitrale vient apaiser l'atmosphère tendue entre les compagnies d'assurances et la CIMR et, ce en adoptant plusieurs mesures visant à améliorer la santé financière de l'organisme et ainsi garantir sa pérennité.
- 2003 est l'année marquante, concernant l'accord CIMR-compagnies d'assurances. A partir de cette année, la CIMR a récupérée la part salariale jusqu'aujourd'hui, qui a été confiée aux compagnies d'assurance. Désormais, les compagnies d'assurances ne gèrent que le stock.

2.2.2 Contexte du travail

Comme cité ci-dessus, à partir de 2003, la CIMR gère la part salariale des contributions de ces affiliés. Une part gérée auparavant par les compagnies d'assurance. Ces dernières n'ont, désormais, qu'un stock constitué des cotisations salariales reçues avant 31 Décembre 2002(mode run off). Les Compagnies d'assurances doivent constituer des réserves réglementaires. Pour ce faire, connaître la loi de sortie des assurés est primordial. Traiter tous les assurés de la même manière ne peut qu'induire en erreurs le calcul de la provision.

En effet, considérer que tous les assurés feront une sortie en *rente viagère* induit à la constitution des réserves qui seront plus importantes que celles constituées dans le cas d'une sortie en *capital*, d'où le risque de surestimation ou de sous-estimation des engagements.

Alors, l'identification des profils de chaque assuré s'avère ainsi importante pour une bonne évaluation des engagements de l'assureur tout en se basant sur les caractéristiques de chaque assuré.

2.2.3 Analyse de la loi interne de la CIMR

La réglementation interne exige un certain nombre de règles concernant la sortie de l'assuré à la retraite. L'adhésion au contrat CIMR permet à un affilié de jouir de tous les droits de mise en retraite.

Types de sorties

- **Sortie pour limite d'âge** : L'âge normal de retraite au sein de la CIMR est 60 ans. A cet âge tout affilié, a droit, selon son choix, à un capital ou à une pension. Toutefois la condition d'au moins 5 ans de contributions étant exigée ainsi que la cessation d'activité dans l'entreprise adhérente afin de bénéficier de sa retraite.
- **Sortie par anticipation** : Tout retraité possédant dans son compte 5 années de cotisations a droit à une retraite anticipée et, ce à partir de l'âge de 50 ans. Dans ce cas, le nombre de points initialement acquis pour l'âge à la retraite à 60 ans est multiplié par un coefficient d'abattement qui réduit le nombre de points pour la retraite anticipée.
- **Rachat des droits** : si un affilié dresse une demande de rachat de ces droits avant l'âge 60 ans, il recevra la totalité de ses cotisations au titre de la part salariale. Toutefois la part patronale est perdue.
- **Cas de la prorogation** : tout affilié, a droit de proroger sa retraite à une date ultérieure à l'âge de sortie normale en retraite (60 ans). Le nombre de points sera affecté d'un coefficient d'augmentation qui sera fonction de l'âge de l'affilié.
- **Cas de rachat de retraite anticipé**: Si le nombre de points d'un participant lors de la liquidation de sa retraite est inférieur à 150 points, il recevra automatiquement de la CIMR un paiement unique de sa retraite. Si le nombre de points de ce participant est compris entre 200 et 150 points il peut demander la liquidation de son paiement en une seule fois sous réserve d'accord de la CIMR.
- **Cas de radiation d'une entreprise adhérente** : Il existe deux motifs importants :
 - ✓ **Démission** : Dans ce cas tous les participants de cette entreprise perdront leur droit à la part patronale. Toutefois, ils garderont le bénéfice de leur part salariale.

- ✓ **Cessation d'activité ou déclaration de faillite** : Les participants à l'entreprise concernée garderont leur droit à la pension au titre de leurs années de services.
- **Cas de décès avant l'entrée en jouissance de la retraite (du point de vue des compagnies d'assurances)** : Deux cas se présentent :
 - ✓ **Cas de participant célibataire** : Dans ce cas, Le bénéficiaire recevra l'ensemble de ces cotisations nettes de taxes et sans addition d'intérêts à la personne désignée comme bénéficiaire.
 - ✓ **Cas des ayants droit** : S'il y a obligation de servir des pensions de réversion ou d'orphelin, la compagnie d'assurance en charge versera ces cotisations à la CIMR qui se chargera de les servir aux ayants droit sous forme de pension.

NB : le contrat du régime commun par capitalisation est un contrat garantissant un capital retraite différé avec option de contre-assurance ; tout décès avant l'âge d'entrée en jouissance de ce capital entraîne automatiquement la récupération des cotisations versées nettes d'intérêts.

- Cas de décès après l'entrée en jouissance de la retraite : plusieurs cas sont possibles selon le type des ayants droits :
 - ✓ Si conjoint seul : il a droit à une pension de 50% de la pension du défunt. L'entrée en jouissance de cette pension est à partir de l'âge normal de retraite pour un conjoint homme et 10 ans plutôt si c'est une femme. Toutefois, le conjoint survivant peut anticiper cette pension jusqu'à 10 ans plutôt.
 - ✓ Si conjoint avec enfant mineur à charge issu du mariage : il a droit à une pension de 50% de la pension normale avec entrée en jouissance immédiat de cette pension (le lendemain du décès) et ce, jusqu'à majorité du dernier enfant mineur à charge. Passée cette date, le traitement du conjoint sera conforme au point ci-dessus.

NB : la condition de deux ans de mariage, avant le décès du participant, doit être vérifiée sauf dans le cas de décès accidentel.

- ✓ **Si orphelin** : il a droit à 20% de la pension de la retraite normale du parent allocataire décédé. Cette pension est versée jusqu'à l'âge de majorité du dernier orphelin. Dans le cas où ce dernier est dans une situation d'invalidité dépassant 66%, alors il a droit de maintenir cette allocation.
- Il existe, toutefois, une **option de retraite à l'âge de 55 ans sans anticipation** : Il faut tout d'abord présenter une demande de retraite à 55 ans sans anticipation. Puis choisir l'une des deux options suivantes :
 - ✓ **Majoration des cotisations** : Lors du choix de l'option de la retraite à 55 ans, toutes les cotisations futures seront majorées de 38,89%. Au moment de 55 ans, pour liquider sa retraite, On appliquera la règle de l'anticipation de la retraite pour tous les points acquis avant la date d'adoption de la retraite à 55 ans. Pour les points acquis après cette date, ils seront intégrés complètement dans le calcul de la pension.
 - ✓ **Rachat de points** : Le rachat de points dans le cadre de cette option tout en bénéficiant de réductions fiscales, permet à l'assuré de bénéficier de sa pension totalement sans réduction. Toutefois, le départ définitif de la société ne permettra à l'assuré de récupérer que ces cotisations salariales et la perte de tous ces adroits à la cotisation patronale ainsi que la majoration payée dans le cadre de la retraite à 55 ans.

Types de liquidation de la retraite

La liquidation des droits d'un assuré ou de ces ayants droits peut être établie sous deux formes :

- **Option capital** : cette option permet à l'assuré ou son ayant droit de liquider la partie salariale gérée par capitalisation sous forme de capital. Tandis que la part patronale est versée sous forme de pensions.
- **Option de rentes** : dans ce cas, la part patronale ainsi que salariale sont versées sous forme de rentes.

2.3 L'ASSURANCE RETRAITE SELON LA VISION D'AXA ASSURANCE MAROC

C'est une filiale marocaine du groupe français AXA spécialisée dans l'assurance créée par la fusion d'AXA Al Aman et la compagnie africaine d'assurances. Elle joue un rôle prépondérant dans le marché marocain et propose une panoplie de produits parmi lesquels on a le produit épargne retraite.

Depuis 1949, après avoir signé la convention de gestion commune entre la CIMR et certains organismes d'assurance, AXA Assurance gère une part importante du portefeuille de la CIMR dans le cadre d'un régime commun de retraite par capitalisation, où elle gère les cotisations salariales en capitalisation et donne le choix aux bénéficiaires de liquider leur retraites soit en rente viagère ou en capital.

En effet, la pension d'un assuré est composée d'une pension versée par la CIMR en contrepartie des contributions patronales et une autre versée par AXA en contrepartie des cotisations salariales.

A cet effet, la Caisse constate à chaque échéance trimestrielle, le montant exact de la pension et la part de cette pension déjà constituée par le régime commun. Elle fournit, alors, les prestations nécessaires pour compléter la pension constituée afin de l'amener au montant de la retraite de base.

La pension servie par la CIMR est égale au montant de la retraite de base au jour de la liquidation diminué du montant de la retraite qui aurait été constituée par les versements inscrits au compte individuel de capitalisation de l'assuré.

L'année de 2003 a été marquée par un avènement historique de nouvelles mesures, quand la CIMR a adopté une réforme suite à des contraintes actuarielles et financières.

Cette réforme a été traduite par la reprise de la gestion en répartition de la part salariale confiée aux compagnies d'assurances et gérée par capitalisation; ce qui implique que les compagnies d'assurances (AXA) ne recevront plus de cotisations à partir de cette date et se contentent de gérer la réserve des cumuls de cotisations dans son portefeuille.

C'est dans cette optique qu'intervient la nécessité de la réévaluation des engagements (provisions et réserves) d'AXA vis-à-vis de ces assurés.

CONCLUSION

Le Maroc dispose de plusieurs caisses de retraite et de compagnies d'assurances qui constituent son système de sécurité sociale. Les modes de fonctionnements de ces organismes sont différents selon les produits garantis. La CIMR, quant à elle, garantit des produits de base et des produits de retraite complémentaire. C'est un régime facultatif qui est géré par capitalisation et par répartition (mixte). Avant 2003, les cotisations des affiliés de la CIMR était gérées en deux parts, une part dite contribution patronales gérées par la CIMR en répartition et une autre dite cotisation salariale gérée en capitalisation par les compagnies d'assurances, dans le cadre d'un régime commun en capitalisation.

La CIMR, est géré par une loi interne qui définit les modalités pour qu'un assuré bénéficie de sa retraite. En effets, l'âge légal de la retraite au sein de la CIMR est de 60 ans. Toutefois, l'assuré a le choix entre plusieurs options pour liquider sa retraite avant ou après l'âge de 60 ans. De plus, la caisse garantit les droits des conjoints et des enfants mineurs et invalides.

Le portefeuille CIMR est d'une importance capital pour AXA Assurance Maroc, puisque les engagements vis-à-vis des assurés de cette compagnie au titre du régime commun en capitalisation sont considérablement importants.

Il est toutefois nécessaire lors de l'évaluation des engagements d'AXA de prendre en considération chaque profil à part, puisque chacun optera pour un type de sortie qui lui est propre et qui n'est pas nécessairement le même que celui d'un autre individu avec les même caractéristiques.

La prise en considération des types de sortie des assurés passe nécessairement par l'identification de ces sorties. La loi interne de la CIMR qui correspond à la charte directrice de la gestion de la CIMR présente une panoplie d'options de sortie qui peuvent être résumées en quatre grandes options : le rachat anticipé de la retraite, la contre assurance, la rente viagère et le capital.

Chapitre 2 : Certification du système interne d'AXA

INTRODUCTION

Dans le système d'information d'AXA, est implémenté tout un ensemble de formules issues des techniques de base du régime commun. Certifier le calcul implémenté est devenu une nécessité afin de mieux appréhender le fonctionnement du système.

Pour ce faire, nous allons commencer par calculer les cotisations, les rentes et les provisions mathématiques suivant les tarifs mis en vigueur par la réglementation.

Chaque année, les versements de l'assuré que de la contractante (employeur), sont entièrement appliqués à la constitution d'une rente viagère différée avec contre-assurance- versements réservés avec échéance à 60 ans.

Le montant de la rente acquise chaque année est déterminé par l'application du tarif légal en vigueur, lors du versement de la cotisation.

1. CADRE LEGAL POUR LES PARAMETRES DE CALCUL

1.1 TARIFS

Depuis sa création, la CIMR a connu plusieurs réformes techniques qui se sont traduites par l'introduction de trois tarifs sur les périodes de réformes.

Un tarif est défini par l'utilisation d'une table de mortalité, un taux pour frais d'acquisition, pour frais de gestion et pour frais de service de la rente et un taux minimum garanti (TMG). Durant ces périodes de réformes, la CIMR a changé un ou plusieurs paramètres, induisant ainsi, un changement du tarif correspondant.

En effet, depuis l'année 1949, le calcul de la rente acquise se fait par le biais de la table RF 46-49 avec un TMG de 3.5%, toutefois les chargements nous ont resté ignorés.

En 1986, après l'introduction de la participation aux bénéficiaires, il y a eu un changement du tarif, désormais le tarif en vigueur étant la PF 60-64, 4.5%, chargement d'acquisition 3%, chargement du service de la rente étant de l'ordre de 3% et les frais de gestion sont de l'ordre de 0.15%.

Enfin, en 1992, suite à l'arbitrage accompli entre les compagnies d'assurances et la CIMR, il y a eu un changement dans les frais d'acquisition qui se sont passés de 3% à 1%.

1.2 PROVISIONNEMENT

D'après la Circulaire de la direction des assurances et de la prévoyance sociale (DAPS) du 13 février 2006 relative à la méthode d'étalement de l'effet du changement des bases techniques en assurances « vie » on a :

« En ce qui concerne les rentes en constitution, les provisions mathématiques sont calculées en adoptant la démarche suivante :

Le contrat sera supposé garantir un capital différé, obtenu par la transformation des rentes acquises sur la base du tarif en usage au moment du paiement des primes, soit :

- Jusqu'au 31 décembre 1985 : table RF 46-49, taux 3,5%
- A partir du 1er janvier 1986 : table PF 60-64, taux 4,5%

Le calcul des provisions mathématiques se fait par l'actualisation de ce capital à la date d'inventaire sur la base de la table de mortalité stipulée par l'article 15 de l'arrêté réglementaire, à savoir : TV 88-90.

En ce qui concerne les rentes en service, le calcul des provisions mathématiques doit se faire sur la base de la nouvelle table de mortalité, à savoir : TV 88-90. ».

2. METHODOLOGIE

2.1 CALCUL DES COTISATIONS ET DES RENTES

Le contrat du régime commun, a pour objet la constitution (avec contre-assurance des cotisations) et le service d'une retraite par capitalisation, moyennant des cotisations périodiques considérées comme une succession de primes uniques chacune procurant un montant de rente différé à l'âge de 60 ans.

2.1.1 Cotisations

La cotisation annuelle (CA) est définie par :

$$CA_t = TC_t * S_t$$

Avec CA_t : Cotisation de l'année t

TC_t : Taux de cotisation de l'année t

S_t : Salaire de l'année t

Etant donné un assuré d'âge x à l'époque 0, l'assureur s'engage à lui verser une rente à partir de l'entrée en jouissance de la retraite. Le calcul de cette dernière se fait par la transformation de la cotisation de l'année en une rente correspondant à la même année et qui sera versée à partir de 60 ans, cette transformation se fait via un coefficient à terme. La rente finale acquise est équivalente à la somme de toutes les rentes correspondant aux années de cotisation.

2.1.2 Calcul du coefficient à terme :

Il s'agit du coefficient de transformation d'une prime de 100 dirhams en rente différée à l'âge de 60 ans et correspondant à un âge x.

$$CAT_{x+k} = \frac{100 * (0.99 - {}_{n-k}A_{x+k})}{1,03 * ({}_{n-k}E_{x+k} + 0.0015 * {}_{n-k}a_{x+k}) * (a_{60} + 0,375)}$$

Avec : x : âge de l'assuré

k : période de versement des rentes

$${}_nA_x = \frac{M_x - M_{x+n}}{D_x} : \text{Assurance décès temporaire de n année.}$$

$${}_kE_x = \frac{D_{x+k}}{D_x} : \text{Capital différé en cas de vie payable après k années.}$$

$${}_na_x = \frac{N_{x+1} - N_{x+n+1}}{D_x} : \text{Annuité viagère temporaire à terme échu sans différé.}$$

$$a_{60} = \frac{N_{61}}{D_{60}} : \text{Rente viagère payable à termes échus (60 ans)}$$

Où $D_x = l_x * v^x$: Première commutation en cas de vie.

$$N_x = \sum_{k=0}^{w-x} D_{x+k} : \text{Deuxième commutation vie}$$

Le taux d'actualisation est de 4.5%

2.1.3 Rente acquise

Les salariés, grâce aux cotisations annuelles qu'ils versent, acquièrent une rente

$$Ra_n = \sum_{k=1}^{k=n} R_k : \text{Rente acquise à l'aide du cumul des cotisations trimestrielles à la fin}$$

de l'année n.

$$\text{Et } R_k = CA_k * CAT_{x+k} : \text{Rente acquise à l'aide de la cotisation de la période k.}$$

CA_k : Cotisation annuelle à la période k et à l'âge x+k.

2.1.4 Rente réversible

La rente réversible R_r pour une fraction λ sur la tête du conjoint :

Si la réversion est au taux $\lambda=50\%$, la rente est réduite à 50% de la rente initiale au décès de la tête de l'assuré par égalisation des capitaux constitutifs de ces deux rentes :

$$R_a * a_{60} = R_r * (a_{60} + \lambda * (a_y - a_{xy}))$$

y : étant l'âge atteint par le conjoint à la date de demande d'entrée en jouissance de la retraite de l'affilié.

a_x : annuité viagère sur la tête de l'assuré x

a_y : annuité viagère sur la tête du conjoint y

$$a_{xy} = \sum_{k=1}^n v^{k+x} \cdot \frac{l_{x+k} \cdot l_{y+k}}{l_x \cdot l_y} : \text{Annuité viagère sur les deux têtes x et y}$$

2.2 CALCUL DES COEFFICIENTS D'ANTICIPATION ET DE PROROGATION

2.2.1 Anticipation :

En égalisant le capital constitutif à l'âge de la retraite normale (60 ans) avec le capital constitutif à l'âge x, nous écrivons la formule suivante :

$$R_x * (a_x + 0,375) = Ra * ({}_k E_x + CFGC * {}_k a_x) * (a_{60} + 0,375)$$

Avec : R_a = rente anticipée ;

R_x : rente acquise.

Le coefficient d'anticipation, de k ans, appliqué à un assuré d'âge x, tel que $x+k=60$, est donné par la formule :

$$C_x = \frac{({}_k E_x + \text{CFG.C. } {}_k a_x) \times (a_{60} + 0,375)}{a_x + 0,375}$$

2.2.2 Prorogation :

De manière analogue, en rapportant le capital retraite constitués à l'âge x à celui constitué pour l'âge de 60 ans, nous pouvons écrire:

$$R_x * ({}_k E_{60} + \text{CFG.C.} * {}_k a_{60}) * (a_x + 0,375) = R_a * (a_{60} + 0,375)$$

Le coefficient de prorogation, qui est un coefficient de revalorisation de la rente acquise à l'âge 60 ans, est donné par la formule :

$$C_x = \frac{(a_{60} + 0,375)}{({}_k E_{60} + \text{CFG.C.} * {}_k a_{60}) * (a_x + 0,375)}$$

Nous comparons les coefficients de prorogation calculés par la formule ci-dessus et ceux données par la CIMR. Toutefois, les résultats de cette comparaison ne seront pas communiqués.

2.3 CALCUL DES RESERVES MATHEMATIQUES

2.3.1 Phase de constitution des rentes

Pendant cette phase, la réserve mathématique est égale à la somme de la réserve mathématique relative à la rente acquise R_j et de la réserve mathématique de la contre-assurance :

$$V = Vr + Vca$$

a) Réserve mathématique de la rente acquise

Elle est égale au taux de réserve mathématique à prime unique pour un dirham de différé à 60 ans, multiplié par la rente acquise.

$$V_{rk} = {}_{n-k}\pi_{x+k} * Ra_k$$

Avec : ${}_{n-k}\pi_{x+k} = (FSR * (a_{60} + 0,375)) * ({}_{n-k}E_{x+k} + CFG * {}_{n-k}a_{x+k})$

b) Réserve mathématique de la contre-assurance

Cette réserve est égale au cumul des cotisations commerciales, multiplié par le taux de réserves à la prime unique pure, soit :

$$Vca_n = \sum_{t=0}^{t=k} \sum_{j=1}^{j=4} (CT_t)_j * {}_{n-k}A_{x+k}, \text{ terme calculé avec la TV88-90,}$$

4.5% comme TMG.

c) Rachat

La valeur de rachat, après k années écoulées depuis l'affiliation au contrat CIMR chez AXA, est égale au cumul des cotisations commerciales nettes d'intérêts :

$$VR_k = \left(\sum_{t=0}^{t=k} \left(\sum_{j=1}^{j=4} (CT_t)_j \right) \right)$$

2.3.2 Phase de service de la retraite

a) Réserve mathématique sans réversion

La réserve mathématique de la rente en service k années après l'entrée en jouissance de la retraite est égale à :

$$V_k = 1,03 * (a_{60+k} + 0,375) * R$$

b) Réserve mathématique avec réversion

La réserve mathématique de réversion pour un assurés de x années et son conjoint qui est âgé de y années est donnée par la formule suivante :

$$V_k = 1,03 * (a_x + \lambda * (a_y - a_{xy}) + 0,375) * R$$

Avec R : rente en service

λ (50%) : coefficient de réversion de la rente au profit du conjoint de l'assuré décédé.

CONCLUSION

L'étape de certification des méthodes de calculs qu'utilise AXA Assurance Maroc, est primordiale afin de cerner le côté technique du régime de la CIMR. D'autre part, cela nous a permis de vérifier l'exactitude des calculs implémentés dans le système d'information d'AXA. Ce qui nous permet de se prononcer sur les engagements que la compagnie doit honorer à l'égard des assurés de la CIMR.

A cet effet, nous avons commencé par recalculer les rentes acquises pour chacun des assurés en utilisant les différents tarifs en vigueur. Nous avons comparé ces résultats avec les calculs implémentés au sein du système d'information. Par la suite, nous avons vérifié l'implémentation de la prorogation ainsi que de l'anticipation dans le système pour comprendre leur fonctionnement. En fin, nous avons certifié le calcul des réserves mathématiques, que ce soit pour le portefeuille des contrats en cours ou pour les assurés ayant choisi l'option rente viagère.

Chapitre 3 : Modélisation de la loi de sortie

INTRODUCTION

Au titre du portefeuille de la CIMR, AXA nous a fourni des données sur l'ensemble des assurés, qui se répartissent sur quatre grandes classes représentant chacune une option de liquidation de la retraite. De plus, ces données contiennent plusieurs catégories qu'il faudra prendre en considération lors de l'évaluation des engagements.

Pour ce faire et afin de mieux évaluer les risques probables liés à l'hétérogénéité de la population, les compagnies d'assurances utilisent des méthodes statistiques spécifiques et puissantes de segmentation (scoring).

Dans l'ouvrage « Data mining et statistique décisionnelle » [Stéphane Tufféry], la notion de « segmentation » ou classement est définie comme « l'opération de placer chaque individu de la population étudiée dans une classe, parmi plusieurs classes prédéfinies, en fonction des caractéristiques de l'individu indiquées comme variables explicatives. L'affectation à une classe à partir des caractéristiques explicatives se fait généralement par une formule, un algorithme ou un ensemble de règles, qui constitue un modèle et qu'il faut découvrir. »

Le principe de notre approche est de réaliser un classement construit sur le comportement historique de la population des assurés de notre portefeuille pour lequel le comportement à la liquidation est connu. Nous allons alors tenter d'approcher ce comportement et de le comprendre à l'aide de règles statistiques et de modèles économétriques de telle façon à faire des prévisions sur les assurés ayants des contrats en cours (actifs).

Nous allons tout d'abord commencer par une étude descriptive des données, puis appliquer notre modèle économétrique afin de faire ressortir les hypothèses nécessaires pour la fin de notre travail. En fin on calcule les provisions mathématiques des individus précités d'après la réglementation en vigueur.

1. PRESENTATION ET PREPARATION DE LA BASE DE DONNEES

Cette partie consiste à purifier la base de données de toutes les erreurs possibles et d'éliminer les doublures causées par la présence d'un individu plusieurs fois dans une même table ainsi que l'agrégation des comptes si un assuré en a plusieurs.

Cette phase de notre travail s'est avérée très difficile en raison d'absence d'une clé unique, identifiant chaque assuré et qui lui est propre. D'autre part la présence de plusieurs tables autres que la table de production (historique) qui est censée contenir toute l'information qui se rapporte à ce portefeuille, nous a posé dans une situation de « step back » et de test ; des tests de nature à comparer les comptes dans toutes les tables dérivées de la table mère qui est la table VPCRCI(table de production ou historique).

La table VPCRCI possède plusieurs identifiants. Un assuré peut être identifié par le numéro de sa police, le numéro du client, le numéro d'affiliation ainsi qu'un numéro de la CIMR ; ce dernier était censé être la clé de notre table mais il n'est pas unique pour chaque assuré. Or les autres tables ne disposent pas de toutes ces informations pour identifier un assuré. La table des affiliés « vpaffil » ne contient pas les numéros de police. Or la présence d'un numéro de police ainsi qu'un numéro d'affiliation et pareillement pour un autre client conduit à la réécriture d'un assuré plusieurs fois car un assuré peut changer son emploi ainsi que de police et numéro d'affilié. Cela conduit à ce que cet assuré possède deux comptes.

Le numéro de matricule de la CIMR est normalement censé donner à chaque affilié le jour de son inscription aux registre de la CIMR à un identifiant qui permettra de le distinguer quoiqu'il change de client ou de police. Dans les tables cette hypothèse s'est avérée fausse.

En effet, un numéro de CIMR peut être commun entre plusieurs assurés. D'où la recherche d'une autre clef était indispensable.

Nous avons essayé dans un premier temps de faire une jointure avec la table des sinistres et la table des affiliés via les trois champs communs : numéro de client, numéro d'affilié et le numéro de la CIMR.

Nous disposons désormais des informations personnelles de chaque client. Nous pouvons assumer que l'identification d'un client passe par son nom, prénom, sa date de

naissance ainsi que le numéro de la CIMR. Celle-ci constitue alors notre clef pour éliminer les doublons.

Nous procédons de la même manière afin de rendre la base de données plus manipulable que ce soit pour les actifs, une table issue de la table «vwrcrci », ou pour l'historique, une table issue de la table « vssini ».

2. MISE EN PRATIQUE DES TECHNIQUES STATISTIQUES

Les variables dans les données utilisées de notre cas sont distinctes (quantitatives et qualitatives) et présentent donc un mélange hétérogène. A cet effet et avant d'appliquer une technique ou une méthode de classement, il est préférable de comparer les différentes techniques de classement et de prédiction en fonction de trois qualités essentielles attendues des meilleurs outils :

- ✚ L'absence d'hypothèses restrictives sur le problème à résoudre.
- ✚ La capacité de traiter toujours exhaustivement les données en un temps raisonnable.
- ✚ La possibilité de manier des données manquantes et de types hétérogènes, numériques ou non (s'agissant des variables explicatives pour les techniques de classement et prédiction). (Stéphane Tufféry, Page 90, Data mining et statistiques décisionnelles)

Tableau 1 : techniques statistiques de prédiction

<i>Techniques</i>	<i>Absence d'hypothèses sur le problème à résoudre.</i>	<i>Traitement exhaustif des bases de données</i>	<i>Traitement des données hétérogènes ou lacunaires</i>
Arbres de décision	Comme la classification hiérarchique (sorte « d'arbre à l'envers »)	Non (mais moins vite limité que la classification hiérarchique)	Certains arbres comme CHAID doivent discrétiser les variables continues
Réseaux de neurones perceptrons	Oui (mais il faut fixer le nombre de neurones cachés)	Non (pas d'apprentissage sur plusieurs centaines de variables)	Les variables non incluses dans [0, 1] doivent être transformées
Réseaux à fonction radiale de base	Comme des perceptrons	Oui	Les variables non incluses dans [0, 1] doivent être transformées
Analyse discriminante	Non (relations linéaires entre les variables et hypothèses sur les lois conditionnelles X_i/Y)	Oui	Variables numériques et sans valeurs manquantes.
Analyse discriminante sur coordonnées factorielles d'une ACM (méthode DISQUAL)	Oui (permet de s'affranchir largement des hypothèses sur les lois conditionnelles X_i/Y)	Oui	Oui (les valeurs manquantes sont traitées comme des valeurs à part entière)
Régression linéaire, régression PLS	Non (relations linéaires entre les variables + autres hypothèses)	Oui	Variables numériques et sans valeurs manquantes
Régression logistique, modèle linéaire généralisé	Oui	Oui	Oui (découper en classes les variables continues avec des valeurs manquantes)

Source : Tableau de Classement et prédiction. (Data mining et statistiques décisionnelles, Stéphane Tufféry)

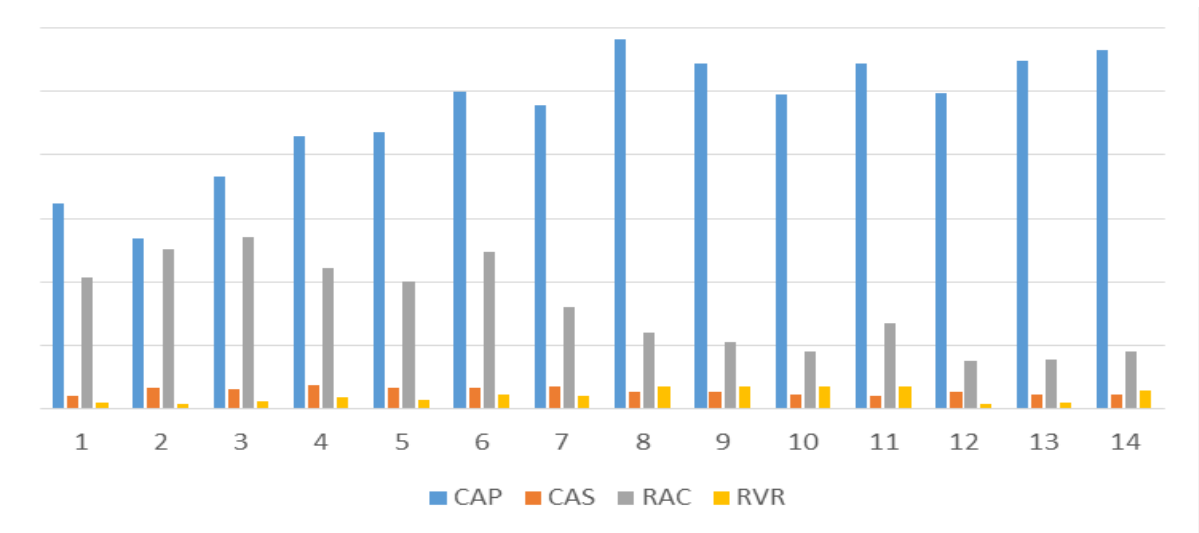
D'après le comparatif des techniques présentées ci-dessus, on déduit que seules les méthodes de la régression logistique (binaire ou multinomiale) et celle de l'arbre de décision conviennent dans notre cas pour la prédiction de la variable dépendante (cible).

2.1 DESCRIPTION DU PORTEFEUILLE

Nous disposons essentiellement de deux tables, une table de l'historique, c'est-à-dire les assurés qui ont quitté le portefeuille (retraités) avant la fin de 2013 et une autre pour les contrats en cours (actifs). Ces tables contiennent toutes les informations sur les assurés. Que ce soit la variable type de sortie qui contient quatre modalités : Rente viagère (RVR), Capital (CAP), Contre-assurance (CAS) et le Rachat (RAC)) ; ou le sexe, le cumul de la rente acquise, le cumul de cotisations, la réserve de la rente, la réserve des cotisations, l'ancienneté et l'âge à l'embauche.

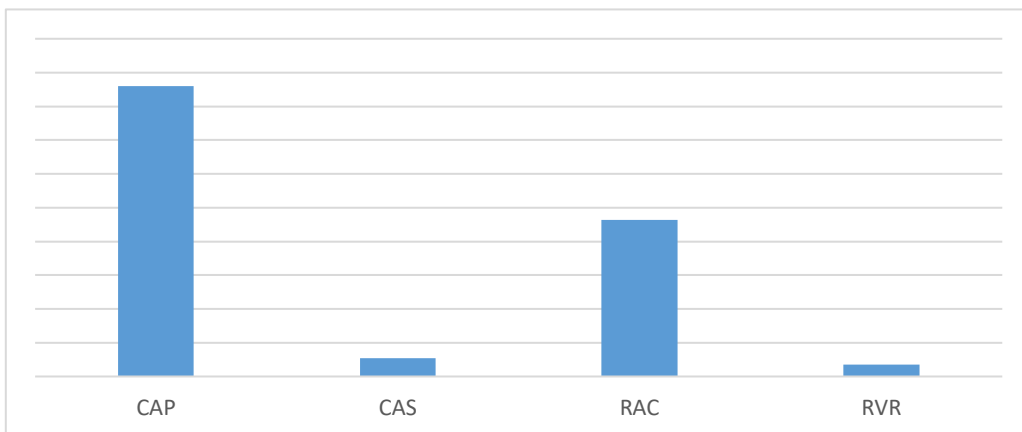
Nous disposons aussi d'autres variables, le nombre d'enfants et le nombre de conjoints, toutefois ces deux variables ne sont pas mises à jour et risquent de fausser notre étude.

Graphique 1 : Répartition de la sortie en fonction des exercices



Nous pouvons remarquer à travers le graphique que l'option la plus prépondérante est l'option Capital, les trois autres options sont marginales en termes de poids. Nous remarquons toutefois, une baisse remarquable de l'option rachat des droits avant l'âge de 50 ans et ce depuis 2000. Cette baisse s'est accompagné d'une augmentation timide de l'option rente viagère. En générale nous avons selon le graphique un nombre de sortie qui reste quasi-stable dans le temps.

Graphique 2 : Répartition des fréquences de la loi de sortie

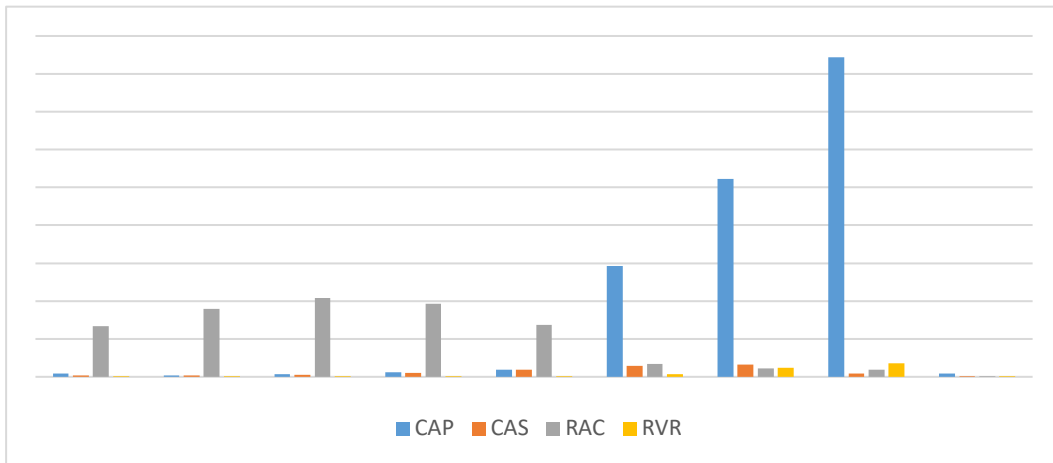


Comme cité ci-dessus, l'option capitale reste l'option qui a le plus de poids suivi de l'option rachat de l'option rente viagère et de l'option contre-assurance.

Nous notons, que l'option capitale est valable lors de la liquidation de la retraite. Chaque assuré a le choix à partir de l'âge de 50 ans de liquider sa retraite en deux options, le capital ou la rente viagère. Ces deux options sont donc complémentaires. Si l'une augmente, l'autre diminue automatiquement.

Nous allons voir plus tard le détail des sorties selon l'âge et ce, pour analyser à quel âge les assurés quittent le régime et en quelle option.

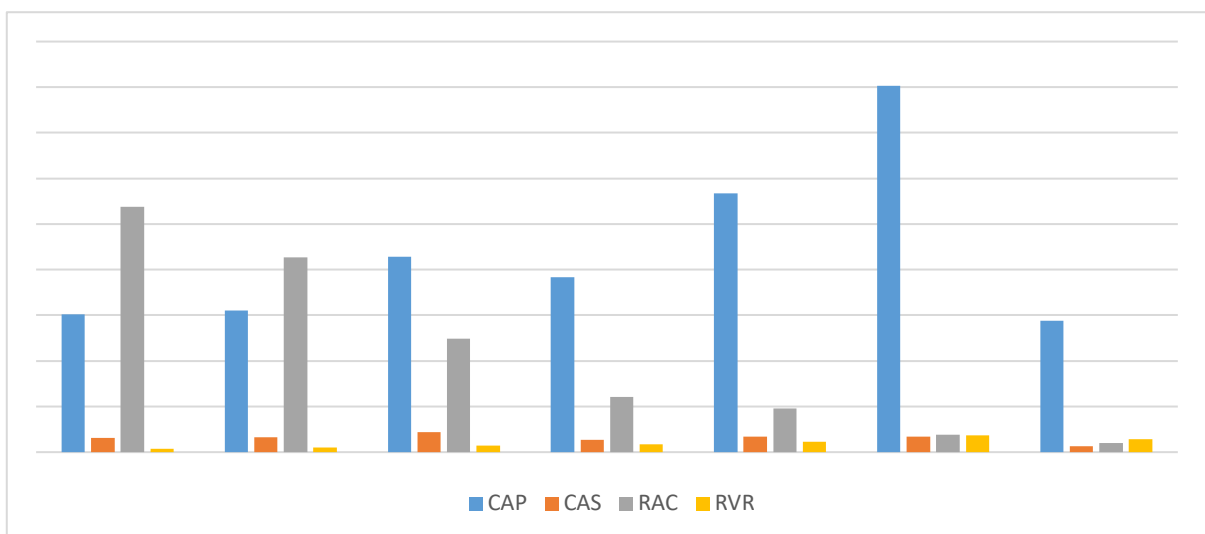
Graphique 3 : Fréquence de sortie en fonction de l'âge et du type de sortie



On peut distinguer à travers le graphique trois grandes périodes, une période d'avant 50 ans, une période de 50 ans à 60 ans et une dernière période des âges supérieurs à 60 ans. La première période est caractérisée par la présence prépondérante de l'option rachat. La seconde est une période où il y a présence majoritaire de l'option capital, cette période correspond à la phase où il y a possibilité d'anticipation de la retraite. La dernière période est une phase où il y'a quasi disparition des sorties.

On peut conclure que la plupart des assurés quittent le régime à l'âge de 60 ans. Pour les autres âges, le poids des sorties reste faible.

Graphique 4 : Répartition de l'historique en fonction de l'ancienneté



On distingue deux grandes tranches, la première jusqu'à une ancienneté de 15 ans et la seconde est supérieure à 15 ans.

La première tranche est caractérisée par l'importance de l'option de rachat et la seconde par l'option de capital. Toutefois pour la première tranche, l'option capital n'est pas du tout négligeable, elle reste légèrement en dessous de l'option rachat.

En revanche à partir de 15 ans d'ancienneté, c'est pratiquement l'option capitale qui est dominante. Cela nous permet de supposer que pour une ancienneté moyenne ou petite, l'assuré a probablement un âge inférieur à 50 ans. Alors que pour les grandes durées d'ancienneté, les assurés ont des âges élevés. D'autre part, le cumul des cotisations croît en fonction de l'ancienneté, puisque cela correspond à une durée de cotisation importante et longue.

Les options, contre-assurance et rente viagère, restent marginales et ne constituent pas un poids important face aux deux autres options.

Tableau 2 : répartition des sorties en fonction du sexe

Sexe\type_sortie	CAP	CAS	RAC	RVR	Total
F	-	108	3299	-	9098
M	74533	-	39950	3170	-
Total	-	-	43249,00	-	-

La population des assurés est une population majoritairement masculine. Cela est dû à la période de souscription au contrat CIMR qui est ancienne. Le sexe n'a aucun effet sur le choix de l'option de sortie, puisque la proportion des hommes par rapport à celle des femmes reste élevée pour tous des sorties.

Nous n'évoquons pas ici les deux autres variables nombre de conjoints et nombre d'enfants des assurés puisque ce sont des informations anciennes et qui ne sont pas mises à jour.

2.2 CADRE THEORIQUE DU MODELE LOGISTIQUE

2.2.1 Le modèle :

L'analyse de la régression est devenue un atout important dans toute étude statistique qui a pour intérêt la description du lien entre une variable réponse et une ou plusieurs

variables explicatives. Le modèle de la régression logistique est l'un des plus utilisés, Il est fréquemment rencontré dans le domaine médical, politique, sciences sociales ...

La différence la plus importante entre les modèles de régression linéaire et logistique est que la variable réponse de ce dernier est binaire (0 ou 1) ou dichotomique (plus de deux modalités). Cette différence se voit dans la forme ainsi que dans les hypothèses des deux modèles. Il suffit uniquement de dépasser la difficulté des variables réponse pour reprendre l'étude en utilisant les techniques de l'analyse de la régression linéaire.

La régression logistique peut être utilisée pour un objectif de description ou de prédiction. C'est aussi une méthode de prédiction supervisée et de discrimination. Ce qui est noté même dans la nature binaire de la variable réponse, celle-là délimite deux classes, les clients font une décision d'acheter ou ne pas acheter un produit, un étudiant peut réussir ou ne pas réussir une année..., chacun des cas possibles fournit deux groupes opposés, l'événement et l'opposé de l'événement.

Puisque la variable réponse est une variable dichotomique ou binaire (0 ou 1) cela va nous empêcher de réaliser des prédictions, alors l'approche usuelle des moindres carrés ordinaires(OLS) qui consiste à minimiser l'erreur autour de la droite de la régression pour une bonne représentation des données ne sera pas efficace. La régression logistique utilise la propriété binomiale de la variable réponse (l'individu appartient au groupe : $Y=1$ avec p , ou, il n'appartient pas au groupe : $Y=0$ avec $1-p$). L'alternative à ce handicap est la technique du maximum de vraisemblance qui consiste à maximiser la probabilité d'appartenance à un groupe sachant les coefficients de la régression.

Dans tout problème de régression, l'intérêt est d'écrire une variable dépendante comme fonction linéaire d'une ou de plusieurs variables explicatives. Dans notre cas (variable réponse binaire), nous allons écrire l'espérance conditionnelle de la variable réponse comme fonction linéaire des variables explicatives.

Désormais, le problème de deux classes de la variable réponse est résolu puisque l'espérance d'une variable de loi Bernoulli est égale à la probabilité de la réalisation de l'événement étudié :

$$Y = \begin{cases} 1 & \text{avec } p \\ 0 & \text{avec } 1-p \end{cases} \quad \text{Où } E(Y/X = x) = P(Y = 1/X = x) = p$$

Or cette écriture n'est pas sans erreur, le côté droit de l'équation de la régression logistique est linéaire et n'est donc pas borné. Ce qui n'est pas le cas du côté gauche de l'équation, puisque c'est une probabilité et est bornée entre 0 et 1. Ceci étant dit, nous avons besoin d'une transformation qui tiendra compte de cette particularité.

La transformation logarithmique ne pourra pas faire l'affaire non plus puisqu'elle sera bornée d'un seul côté (borne supérieure). La transformation logistique remédie parfaitement à ce problème puisqu'elle n'est pas bornée, si la partie linéaire est positive cela signifie que l'individu avec les caractéristiques X sera responsable de l'événement Y=1, le cas contraire est juste.

Le modèle désormais étudié est :

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \sum_{i=1}^p \beta_i * X_i$$

En ce qui concerne les variables explicatives ou variables prédictives, le modèle de régression logistique prend en considération les variables qualitatives ainsi que les variables continues ce qui fait de lui l'alternative ou bien le premier outil de l'analyse discriminante et les études de scoring. Cependant, si les observations sont trop dispersées ou s'il y'en a peu, ce serait difficile de calculer les probabilités $P(Y = 1 / X = x)$ directement. Pour y remédier il serait nécessaire de procéder à un regroupement des valeurs de X.

2.2.2 Estimation des paramètres

La régression logistique en particulier et les méthodes GLM en général, diffèrent de la régression linéaire usuelle du fait que les paramètres ne sont pas estimés via la méthode des moindres carrés ordinaire, mais par la méthode du maximum de vraisemblance. Pour estimer les paramètres, on doit tout d'abord construire la fonction de vraisemblance L. Les paramètres β_i qui maximisent la fonction de vraisemblance sont appelés les estimateurs de maximum de vraisemblance. Cette fonction s'écrit sous la forme :

$$L(\beta, x) = \prod_{i=1}^n p^{y_i}(x_i) * (1 - p(x_i))^{(1-y_i)}$$

Il est mathématiquement simple de travailler avec le logarithme de la fonction de vraisemblance au lieu de la fonction définie initialement. Ce qui donne :

$$\ell(\beta, \mathbf{x}) = \sum_{i=1}^n y_i * \log(p(x_i)) + (1 - y_i) * \log(1 - p(x_i))$$

Pour maximiser la fonction de vraisemblance on dérive la log-vraisemblance par rapport à chacun des β_j et on égalise ces équations à 0. Ce qui nous donne $p+1$ équations à résoudre. Ces équations sont de la forme :

$$\begin{cases} \sum_{i=1}^n (y_i - p(x_i)) * x_j = 0 \\ \sum_{i=1}^n (y_i - p_i) = 0 \end{cases} \quad \text{ou } j = 1, \dots, p$$

L'absence d'une formule analytique fermée est l'un des principaux inconvénients de cette méthode. Pour résoudre les équations il est primordial de faire appel à des méthodes itératives. La méthode de Newton-Raphson étant la meilleure.

Ceci étant fait, il est judicieux de tester la validité des estimations. Une estimation avec une grande variance n'est pas trop robuste et réduit ainsi la chance que la variable correspondante ait un impact significatif sur la variable réponse étudiée.

Nous proposons par la suite des tests pour mesurer la qualité des estimations et ainsi du modèle résultant.

2.2.3 Tests de validité des estimations et du modèle

2.2.3.1 Test de signification des coefficients

Comme en régression linéaire, les tests de signification des estimations peuvent être faits séparément ou collectivement.

La Statistique de Wald pour une variable X_i peut être utilisée pour mesurer la signification de l'estimation d'un paramètre donné, ou encore la contribution d'une variable X_i dans le modèle étudié. Cette estimation est d'autant significative que la Statistique de Wald correspondante est supérieur à 3,84.

L'hypothèse testée dans ce cas $H_0 : \beta_j = 0$; on accepte H_0 si la valeur de la statistique de Wald est supérieure à 3,84, ou si la p-value correspondante est inférieure à 5%.

Ce n'est cependant pas sans défaut, lorsque les coefficients β_j sont grands, le test peut échouer à rejeter l'hypothèse nulle même lorsque ces coefficients sont significativement différents de 0¹. La faiblesse de ce test est expliquée par le fait que la source de l'aléa dans la statistique du test provient de l'estimation du paramètre dans le numérateur et l'estimation de la variance dans le dénominateur.

2.2.3.2 Test de la Déviance

On définit la Déviance d'un modèle à k degrés de liberté (chaque variable quantitative donne 1 degré de liberté et chaque variable qualitative à c modalités donne $c-1$ degrés de liberté) par :

$$D(\beta_k) = -2 * [\log(L(\beta_k)) - \log(L(\beta_{\max}))] = -2 * \log\left(\frac{L(\beta_k)}{L(\beta_{\max})}\right)$$

Avec $L(\beta_{\max})$: est la fonction vraisemblance du modèle saturé(ou modèle à seule la constante).

Dans le cas où la variable cible est une variable binaire, $L(\beta_{\max}) = 1$.

Alors : $D(\beta_k) = -2 * \log(L(\beta_k))$

La déviance peut être utilisée comme un test global pour tester la signification d'un ensemble de variables ou l'ajustement d'un modèle. On peut aussi l'utiliser pour tester la nullité d'un coefficient et ce en calculant la différence :

$$D(\beta_k) - D(\beta_{k+1}) = -2 * [\log(L(\beta_k)) - \log(L(\beta_{k+1}))] ;$$

On considère l'hypothèse nulle $H_0 : \beta_1 = \dots = \beta_l = 0$. Sous cette hypothèse la différence de déviances $D(\beta_k) - D(\beta_{k+1})$ suit une loi de χ^2 (Khi-deux) de l degrés de liberté.

On peut énoncer le test global tel que l'hypothèse nulle : tous les paramètres estimés du modèle sont significativement différents de 0, alors le modèle s'ajuste très bien avec les données. Cela étant dit, sous l'hypothèse H_0 le modèle ajusté suit une loi de χ^2 à $n - p$ degrés de liberté.

¹ Stéphane Tufféry

Ce que nous cherchons à obtenir c'est un modèle avec une Déviance qui soit petite par rapport à celle du modèle saturé. L'ajout d'une variable ne fait que réduire cette mesure, mais la question qu'il faut se poser est ce que cette quantité a diminué suffisamment pour rejeter l'hypothèse H_0 de la nullité du coefficient β_{k+1} correspondant à la variable X_{k+1} . D'autres mesures peuvent être calculées. Celles-là pénalisent encore plus le calcul de la Déviance :

AIC : The Akaike Information Criterion : $= -2 * \log(L(\beta_k)) + 2 * (k + 1)$;

BIC : The Bayesian information Criterion : $= -2 * \log(L(\beta_k)) + (k + 1) * \log(n)$;

Odds Ratio : L'Odds ratio d'une variable explicative mesure l'impact sur la variable réponse résultant d'une variation d'une unité de la variable explicative X.

Si la variable X_i varie d'une unité c'est-à-dire de x à $x+1$, dans ce cas la valeur du $\text{logit}(p(x))$ varie de la quantité β_i , en d'autres termes :

$$OR = \frac{\frac{p(x+1)}{(1-p(x+1))}}{\frac{p(x)}{(1-p(x))}} = \exp(\beta_i)$$

Si une variable est qualitative avec c modalités, on pourra alors calculer $c-1$ Odds ratio. La modalité restante est choisie comme modalité de référence. Pour les variables continues, on n'a pas besoin d'une référence. Mais le problème que peut poser ces variables c'est qu'il se peut que la non-linéarité ne soit pas prise en considération ainsi qu'il y ait une faible robustesse du modèle (selon Stéphane Tufféry).

Si le paramètre β_i estimé est positif alors la variable X influe positivement sur la variable réponse. S'il est négatif alors la variable explicative X influe négativement sur la variable réponse. En fin, si la variable $\beta_i \equiv 0$ alors le $OR = 1$ alors la variable n'a pas d'impact sur la variable réponse, par la suite elle n'est pas significative. Ainsi tout intervalle de confiance de l'Odds Ratio contenant la valeur 1 indique que la variable correspondante n'as pas un apport significatif sur l'explication de la variable réponse Y.

2.2.3.3 La courbe ROC

Avant de présenter une définition de la courbe ROC, il est important de définir deux entités : la spécificité et la sensibilité.

- ✓ **La sensibilité** : c'est la proportion des individus pour lesquels on a prédit la détection de l'événement et pour lesquels on a vraiment détecté l'événement. C'est la proportion des individus correctement diagnostiqués par le test comme sujet de l'événement.
- ✓ **La spécificité** : c'est la proportion des individus pour lesquels on n'a pas vraiment détecté l'événement et qui sont correctement diagnostiqués par le test (correctement prédits).

		La vérité	
		"+"	"-"
le modèle	"+"	a	b
	"-"	c	d

$$\text{Sensibilité} = \frac{a}{a + c} ; \text{Spécificité} = \frac{d}{d + b}$$

La Courbe ROC (Receiver Operating Characteristic) est la représentation de la *Sensibilité* en fonction de la variable $1 - \text{Spécificité}$. Cette dernière mesure représente la proportion des individus positivement prédits par le modèle et pour lesquels on ne détecte pas en réalité l'événement.

Plus la zone sous la courbe (Area Under the ROC) est proche de 1, plus l'ajustement est bon. Autrement dit, plus notre courbe est loin de la première bissectrice et proche du coin gauche du graphique, plus le modèle s'ajuste très bien et on a alors peu d'individus mal classés. Si la courbe correspond à la première bissectrice alors le classement effectué par le modèle est aléatoire.

2.2.3.4 Autres tests statistiques en régression logistiques

Plusieurs tests peuvent être cités à ce stade, quelques-uns ont été abordés précédemment. Pour récapituler, nous présenterons ci-après quelques tests qui peuvent être menés que ce soit pour tester la signification des coefficients ou pour tester le bon ajustement du modèle :

- ✚ test de Wald pour tester la signification des paramètres.
- ✚ Le test de l'Odds Ratio pour tester la signification de la contribution de chacune des variables explicatives du modèle.
- ✚ Le test de la Déviance pour faire des tests locaux sur la signification des variables ou des tests globaux pour tester le bon ajustement du modèle ;
- ✚ Le BIC et l'AIC qui dépendent du nombre d'observations dont nous disposons.
- ✚ Le R^2 de Cox-Snell et le R^2 ajusté.
- ✚ Test de Hosmer-Show qui concerne les proportions théoriques et calculées.
- ✚ Test de Déviance normalisé et la statistique de Pearson.
- ✚ Tests de Concordance : indice de Gini et la courbe ROC.

2.3 ARBRES DE DECISION

2.3.1 Démarche

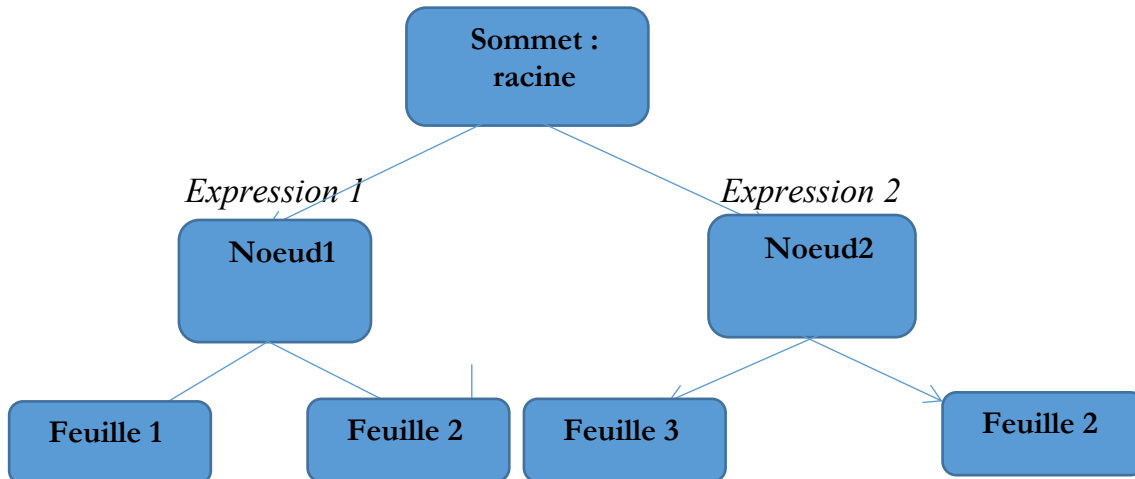
La technique de l'arbre de décision est un modèle basé sur un arbre ou ce qu'on appelle une classification hiérarchique descendante. Elle a un double objectif, elle permet d'une part de répartir les individus d'une population en n classes prédéfinies et d'autre part, permet de prédire l'appartenance de nouveaux individus à l'un des groupes ou classes prédéfinis en fonction de ses caractéristiques.

On commence par choisir la variable qui, par ses modalités, sépare le mieux les individus de chaque classe, de façon à avoir des sous-populations, que l'on appelle nœuds, contenant chacune le plus possible d'individus d'une seule classe, puis on réitère la même opération sur chaque nouveau nœud obtenu jusqu'à ce que la séparation des individus ne soit plus possible (au vu de certains critères dépendants du type d'arbre).

Par construction, les nœuds terminaux (les feuilles) sont tous majoritairement constitués d'individus d'une seule classe avec une forte probabilité, quand il satisfait l'ensemble des règles permettant d'arriver à cette feuille².

² Stéphane Tufféry, Page 184, Data mining et statistiques décisionnelles

2.3.2 Structure d'un arbre de décision



- ✚ sommet ou racine : le premier nœud de l'arbre.
- ✚ feuilles : nœuds terminaux qui ne sont plus segmentés.
- ✚ nœuds-fils : sous-nœuds.
- ✚ expression : chemin entre chaque racine et une feuille.
- ✚ score d'appartenance d'un individu à une classe : pourcentage de ces individus dans la même classe.
- ✚ densité d'un nœud : rapport de son effectif par rapport à l'effectif total de la population.

2.3.3 Types d'algorithmes

Les principaux algorithmes d'arbre sont :

- ✚ CART (classification and Régression Tree), adapté à l'étude de tout type de variables.
- ✚ C5 .0 (de J.R Quinlan), adapté à l'étude de tout type de variables
- ✚ CHAID (Chi-Square Automation Interaction Detection), initialement réservé à l'étude des variables explicatives et cible, discrètes et catégorielles.

Dans notre cas, nous choisissons l'algorithme CHAID, car il est le plus simple en termes de construction d'un arbre de décision.

2.3.4 Étapes de la construction de l'arbre de décision (CHAID)

Deux étapes principales pour la construction de l'arbre de décision (CHAID) à savoir :

- ✚ **Choix de la variable qui sépare le mieux les individus de chaque classe** : Une fois la meilleure séparation trouvée, on l'applique et on réitère jusqu'à ce qu'on trouve un ou plusieurs nœuds-fils et ainsi de suite. On continue les itérations jusqu'à ce qu'aucune séparation ne soit possible selon des critères précis.
- ✚ **Répartition des individus dans les nœuds** : Affecter chaque individu exactement dans une feuille.

2.3.5 Algorithme CHAID

On utilise généralement, le test du χ^2 pour définir la variable la plus significative de chaque nœud et ce, pour des variables explicatives discrètes ou catégorielles. La plupart des logiciels ayant implémenté CHAID (comme SPSS), peuvent traiter, d'une part, les variables explicatives continues, en les discrétisant et d'autre part les valeurs manquantes en les affectant à une seule catégorie.

Le test de χ^2 est utilisé dans les deux étapes successives de division de chaque nœud :

- 1) On utilise pour chaque variable explicative X, le χ^2 pour regrouper ses modalités en les croisant avec les modalités de la variable cible.
Si χ^2 pour la paire de modalités de X n'est pas significatif au seuil choisi, on procède à une fusion des deux modalités en considérant leur résultat comme une nouvelle modalité et ainsi de suite jusqu'à ce que toutes les paires de modalités aient un χ^2 significatif.
- 2) Si toute modalité composée est constituée d'au moins trois modalités initiales, on détermine la division binaire au plus grand χ^2 , on effectue la division au cas où le χ^2 est significatif.
- 3) Après le regroupement des modalités pour chaque variable et le calcul de χ^2 correspondant, CHAID retient la variable dont le χ^2 est le plus significatif, selon un critère prédéfini et dépendant non seulement de la probabilité associée au χ^2 mais aussi éventuellement d'un coefficient correctif dit de Bonferroni.

Si χ^2 est plus significatif que le seuil choisi, alors le nœud est divisible en autant de nœuds-fils qu'il y'a de modalités après changement.

Si χ^2 n'atteint pas le seuil spécifié, le nœud n'est pas divisé³.

2.4 RESULTATS

2.4.1 Problématique

Nous voulons prédire le nombre de sortie dans chacune des options que présente la loi interne de la CIMR et calculer les réserves correspondantes à chacune des options. Pour cela, un modèle logistique a été choisi. Ce modèle nous permettra d'expliquer la variable type de liquidation avec les variables disponibles dans notre base de données à savoir :

- **SICTYP** : variable dépendante qui représente le type de sortie (quatre modalités : RVR, CAP, CAS et RAC).
- **CRMCCO_c** : représentant les tranches du cumul des cotisations de chaque assuré.
- **AFCSEX** : le sexe de l'assuré.
- **Ancienneté** : présente la durée passée par l'assuré dans le système comme étant un affilié.
- **Age_embauche** : c'est l'âge de l'assuré, lors de son embauche.

Nous avons tenté d'expliquer la variable cible qui est le type de sortie par les variables explicatives suivantes : Age_embauche, ancienneté, CRMCCO et AFCSEX. Nous éliminons les variables CRMRE, CRMCO, CRMCRE, nombre d'enfants et nombre de conjoints que ce soit pour raison de corrélation ou d'ancienneté de l'information.

La corrélation à ce stade n'est pas vérifiée par un test puisque chacune des trois variables CRMRE, CRMCO et CRMCRE découlent de la variable CRMCCO.

Pour les insérer dans un modèle de régression logistique, nous allons tout d'abord commencer par recoder certaines des variables explicatives continues sous forme de tranches à savoir la variable CRMCCO. Nous analysons ensuite la distribution de cette variable à l'aide de l'algorithme CHAID sous SPSS. Ce dernier nous donne le nombre de tranches à considérer ainsi que le découpage optimal qui en résulte de telle façon que chacune des tranches soit la plus significative possible. En même temps, nous transformons la variable ancienneté en utilisant le logarithme.

³ Stéphane Tufféry, Page 192, Data mining et statistiques décisionnelles

2.4.2 Modèle logistique pour l'option Capital

2.4.2.1 Le modèle

Nous allons utiliser dans un premier temps le modèle logistique pour l'option capital. La variable cible « CAPITAL » est sous la forme :

$$\text{Capital} = \begin{cases} 1 & \text{si l'option "capital" est choisie avec une probabilité } p \\ 0 & \text{sinon} \end{cases}$$

Le modèle à estimer s'écrit alors sous la forme :

$$\text{logit}(p) = \text{age_embauche} + \log(\text{ancienneté}) + \text{CRMCCO_c}$$

Où p est la probabilité de choisir l'option capitale.

Nous étudions tout d'abord la corrélation entre les variables continues et mesurons l'association entre les variables explicatives à savoir Age_embauche et Log (ancienneté) d'une part et la variable CRMCCO_c d'autre part.

A cet effet, Nous utiliserons la régression logistique afin de mesurer le degré d'association entre la variable qualitative CRMCCO_c et les deux variables explicatives continues age_embauche et log (ancienneté).

2.4.2.2 Corrélation

Nous étudions la corrélation entre les variables explicatives continues à l'aide de la matrice de corrélation de Pearson :

Tableau 3 : la matrice de corrélation de Pearson.

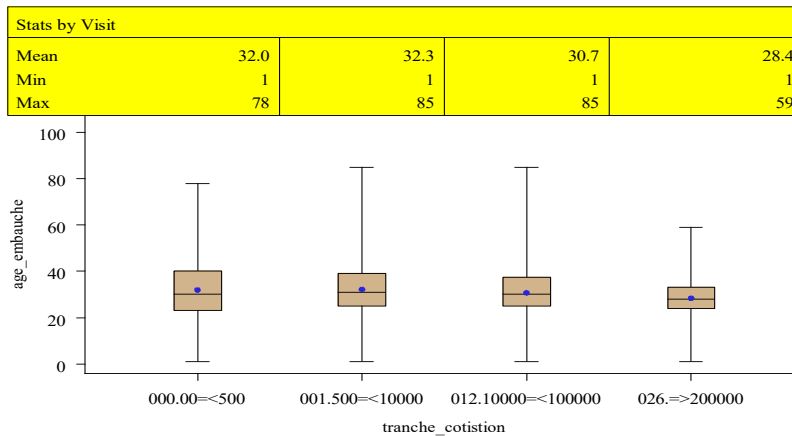
Pearson Correlation Coefficients, N = 84360 Prob > r under H0: Rho=0		
	age_embauche	Log (ancienneté)
age_embauche	1.00000	-0.35750 <.0001
Log (ancienneté)	-0.35750 <.0001	1.00000

Selon la matrice ci-dessus, nous pourrions conclure que la liaison linéaire entre l'âge à l'embauche (age_embauche) et l'ancienneté de l'assuré dans le régime (log(ancienneté)) est négativement faible, puisque le coefficient de corrélation de Pearson est égal à -0.35, qui reste largement inférieur à 1.

Pour avoir une idée sur la liaison entre l'ancienneté et l'âge à l'embauche en fonction du cumul des cotisations des assurés, nous traçons ci-dessus les boîtes à moustaches de chacune des variables continues en fonction de la variable qualitative CRMCCO_c.

Cas de la liaison entre l'âge à l'embauche et le cumul de cotisation :

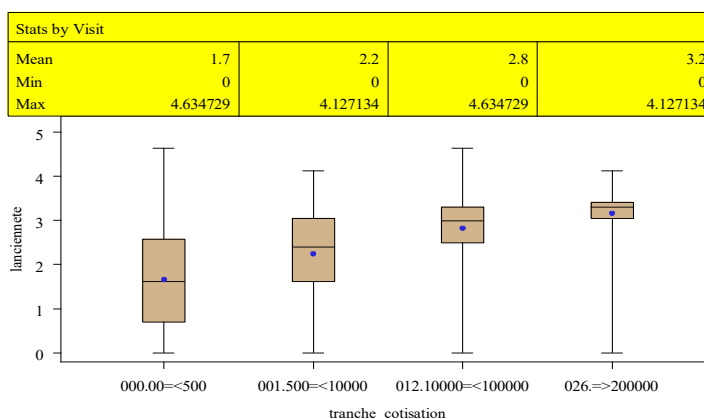
Graphique 5 : Répartition de l'âge à l'embauche en fonction des tranches du cumul de cotisation



Pour le graphique ci-dessus, qui présente l'âge à l'embauche en fonction de la tranche de la cotisation cumulée, nous remarquons que la moyenne est quasi-confondue avec la médiane, ce qui signifie que les distributions de l'âge à l'embauche dans chacune des tranches du cumul de cotisation sont symétriques. D'autant plus, les médianes des différentes tranches du cumul de cotisation sont pratiquement dans le même niveau. Ce qui signifie que l'âge à l'embauche se distribue d'une manière pratiquement indépendante aux tranches du cumul de cotisation. Seul l'écart inter quantile a tendance à se rétrécir si on se déplace vers des tranches supérieures de cumul de cotisation.

Cas de la liaison entre l'ancienneté et le cumul de cotisation :

Graphique 6 : Répartition d'ancienneté en fonction de tranche cotisation



Pour ce graphique, qui présente l'ancienneté en fonction des tranches de cotisation des assurés, il y a lieu de constater que l'écart inter quantile se rétrécit avec le passage d'une tranche inférieure vers une tranche supérieure du cumul de cotisation. Ce qui signifie que la population des assurés devient de plus en plus concentrée, selon le critère de l'ancienneté. D'un autre côté, la moyenne se déplace vers le bas et en dessous de la médiane, ce qui signifie que la distribution devient de plus en plus étalée vers la droite. De ce fait, nous pouvons soupçonner l'existence d'une relation entre l'ancienneté dans le régime et le cumul de cotisation.

Nous allons calculer le coefficient de corrélation pour chacun des deux cas et ce, en décomposant la variance de log (ancienneté) et celle de l'age_embauche.

La formule de décomposition s'écrit de la façon suivante :

$$\sigma^2_y = \frac{1}{n} \sum_{l \geq 1} n_l (\bar{y}_l - \bar{y})^2 + \frac{1}{n} \sum_{l \geq 1} n_l \sigma_l^2 = \sigma_E^2 + \sigma_R^2$$

Y étant la variable quantitative et X la variable qualitative à m groupes.

Avec :

- \bar{y} : Moyenne pondérée par les moyennes de chaque groupe.
- \bar{y}_l : Moyenne du groupe l.
- σ_E^2 : la variance expliquée qui traduit la variabilité inter-groupe.
- σ_R^2 : la variance résiduelle qui traduit la variabilité intra-groupe

Nous calculons alors le coefficient de corrélation à l'aide de la formule suivante :

$$s_{Y/X} = \sqrt{\frac{\sigma_E^2}{\sigma_R^2}}$$

Ce coefficient est compris entre 0 et 1. Plus le coefficient est proche de 1, plus la liaison entre X et Y est forte. Cette liaison est dite parfaite lorsque le coefficient de corrélation est égal à 1.

Ainsi, le coefficient de corrélation pour chaque couple des variables se présente comme suit :

Tableau 4 : Coefficient de corrélation pour chaque couple des variables

paramètres	Log(Ancienne)/CRMCCO_c	Age_embauche/CRMCCO_c
σ^2_E	0,300906978	1,060899875
σ^2_R	0,684280268	110,2610756
$S_{Y/X}$	0,66313068	0,098090305

Selon le tableau ci-dessus, le coefficient de corrélation nous permet de confirmer les conclusions tirées à travers l'analyse des boîtes à moustache : Absence de corrélation entre l'âge à l'embauche et le cumul de cotisation, à l'inverse de la relation entre l'ancienneté et le cumul de cotisation qui reste acceptable. Toutefois, le coefficient de corrélation ne dépasse pas le seuil de 0,8.

En sommes, nous retenons les trois variables (l'âge à l'embauche, l'ancienneté et le cumul des cotisations) comme variables explicatives du modèle logistique que nous avons déroulé sous SAS.

2.4.2.3 Estimation des paramètres

Nous avons exécuté sous SAS la procédure **Proc logistic** afin d'estimer les paramètres du modèle logistique défini auparavant (voir paragraphe 2.4.2.1, page 45) :

Tableau 5 : Résultats du model sous SAS

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	111213.74	80839.610
SC	111223.08	80895.667
-2 Log L	111211.74	80827.610
R-Square	0.3024	
Max-rescaled R-Square	0.4129	

Les statistiques d'ajustement nous montrent que la quantité $-2*(\text{Log L})$ a diminué considérablement lors de l'ajout des Co-variables au modèle saturé. Nous avons toutefois un R^2 qui reste faible.

Tableau 6 : test de signification globale

Type III Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
age_embauche	1	11367.9911	<.0001
l'ancienneté	1	12580.5992	<.0001
CRMCCO_c	3	737.0723	<.0001

Toutes les variables explicatives introduites dans le modèle ont un effet sur la variable expliquée. Cela est illustré par le test du Type III de la sortie de SAS. En effet, une variable a un effet sous l'hypothèse nulle si la statistique de Wald qui lui est propre est inférieure à la valeur théorique de χ_{DF}^2 avec un niveau de signification $\alpha = 0.05$:

Tableau 7 : Estimation des coefficients du modèle de capital

Analysis of Maximum Likelihood Estimates								
Parameter		Capital	DF	Estimâte	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept		Ncapita	1	6.1464	0.0798	5930.7582	<.0001	467.047
age_embauche		Ncapita	1	-0.1122	0.0010	11367.991	<.0001	0.894
Log (ancienneté)		Ncapita	1	-1.5750	0.0140	12580.599	<.0001	0.207
CRMCCO_c	000.00=<500	Ncapita	1	0.6313	0.0547	133.3101	<.0001	1.880
CRMCCO_c	001.500=<10000	Ncapita	1	0.9386	0.0492	364.2279	<.0001	2.556
CRMCCO_c	012.10000=<100000	Ncapita	1	0.4891	0.0485	101.8454	<.0001	1.631

Les coefficients estimés des sont tous significatives au niveau de 0.005, ce qui entraîne que nous rejetons l'hypothèse nulle de la nullité des coefficients du modèle. Pour un individu de référence avec les caractéristiques suivante : age_embauche=0, l'ancienneté=0 et CRMCCO_c= « >=100000 » avec une option capital lors de la sortie, le prédicteur linéaire correspondant est de 6,1464. Nous appellerons cela, score de référence. Si l'âge à l'embauche augmente d'une année, cela entraînera une diminution de la chance de ne pas avoir une sortie capitale par rapport à l'individu de référence de l'ordre de 10%. Dans la même optique, log (ancienneté) réduit à son tour cette chance de sortie capital par rapport à l'individu de référence de 80%. Cela est normal puisqu'un individu qui a commencé à travailler à un âge très jeune ne peut que cotiser plus, ainsi de même pour l'ancienneté dans le régime, plus on est ancien plus la cotisation accumulée est élevée. Cependant, plus un individu passe de la tranche de référence à une autre tranche, plus sa chance de ne pas avoir une sortie capitale augmente.

Tableau 8 : Statistiques d'ajustement du modèle de capital

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	84.1	Somers' D	0.685
Percent Discordant	15.7	Gamma	0.686
Percent Tied	0.2	Tau-a	0.319
Pairs	165955634	c	0.842

La statistique C et le pourcentage de concordance nous donnent une idée sur la bonne classification du modèle. Ils sont de l'ordre de 85%, ce qui signifie que 85% des valeurs observées ont été correctement prédites et seules les 15% de ces observations sont mal prédites

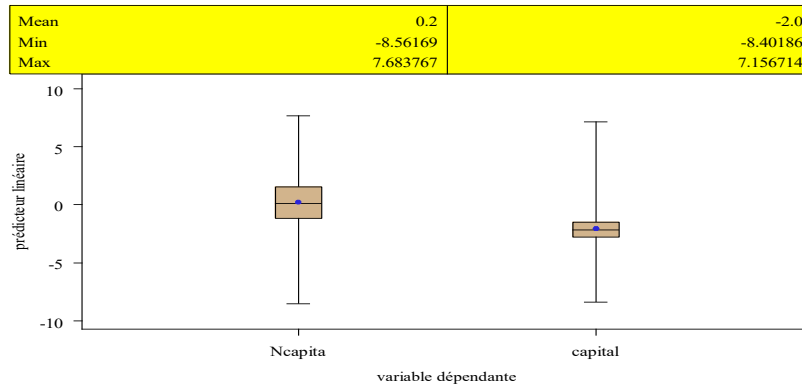
Tableau 9 : Table de confusion pour le modèle du capital

Table of _FROM_ by _INTO_			
FROM (Formatted Value of the Observed Response)	_INTO_ (Formatted Value of the Predicted Response)		
Frequency Percent Row Pct Col Pct			
	Ncapita	capital	Total
Ncapita	18735 22.21 59.96 79.33	12509 14.83 40.04 20.59	31244 37.04
capital	4882 5.79 9.19 20.67	48234 57.18 90.81 79.41	53116 62.96
Total	23617 28.00	60743 72.00	84360 100.00

Le tableau ci-dessus présente en ligne les valeurs observées des sorties à la retraite selon les deux modalités : sortie en capital et sortie en non-capital. En colonne il présente les valeurs prédites des sorties à la retraites par le modèle en faisant une répartition de chaque modalité observée selon les deux modalités (sortie en capital et sortie en non-capital).

Ainsi, 59% des sorties à la retraite en option non-capital sont classés correctement. Autrement dit, 40% de ces assurés sont mal classés par le modèle. De même, 91% des sorties à la retraite qui ont effectivement choisi l'option capital sont bien classées et uniquement près de 9% sont mal classés par le modèle.

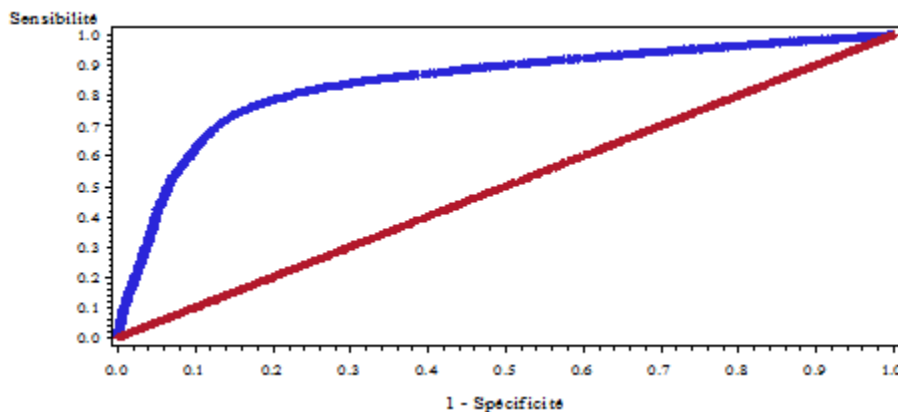
Graphique 7 : Répartition du logit(p) en fonction de l'option capital



Dans cette boîte à moustache, nous avons étudié la distribution du prédicteur linéaire (Logit (P)) en fonction des deux options de sortie : capital et non-capital.

Il est à remarquer que les deux boîtes à moustaches ne sont pas dans le même niveau. Cela signifie que les variables choisies pour expliquer la variable dépendante sont bien discriminantes. La moyenne de la distribution des prédicteurs linéaires pour les assurés ayant choisi l'option capitale est de -2 tandis que celle pour les assurés n'ayant pas choisi cette option est de 0,2. Cela montre effectivement une très grande différence entre les deux groupes. De plus l'écart inter quantile de l'option capital est beaucoup moins faible que celui correspondant au groupe de l'option non capital.

Graphique 8 : La courbe ROC du modèle de l'option capital



La courbe ROC ci-dessus confirme les résultats précités. Ainsi, la surface entre la courbe ROC (en bleu) et la première bissectrice (Rouge) est assez élevée. Ce qui veut dire que l'ajustement du modèle est acceptable.

Nous allons par la suite effectuer la même étude pour l'option Rachat afin de parvenir à prédire correctement les sorties des assurés. Nous allons dresser le modèle, prédire les options et les croiser avec celles déjà calculées pour l'option capitale. Nous nous

intéressons aux assurés qui auront comme option prédite par les deux modèles : Ncapital et Nrachat. Ces assurés sortiront, alors, soit par contre assurance ou par rente viagère.

2.4.1 Modèle de la loi de rachat

2.4.1.1 Le modèle

Le modèle écrit pour l'option rachat est le même que le précédent tel que :

$$\text{rachat} = \begin{cases} 1 & \text{si l'option rachat est choisie avec une probabilité } p \\ 0 & \text{sinon} \end{cases}$$

Ainsi, le modèle logistique s'écrit de la même façon que celui que nous avons construit pour le cas du capital :

$$\text{logit}(p) = \text{age_embauche} + \log(\text{ancienneté}) + \text{CRMCCO_c}$$

2.4.1.2 Estimation des paramètres

Nous exécutons la procédure **Proc Logistic** sous SAS et nous avons obtenu les résultats ci-dessous :

Tableau 10 : statistiques d'ajustement du modèle de rachat

Model Fit Statistics		
Criterion	Intercept Only	Intercept andCovariates
AIC	102937.91	65313.010
SC	102947.25	65369.067
-2 Log L	102935.91	65301.010
R-Square	0.3599	
Max-rescaled R-Square	0.5106	

Il y a lieu de constater que la vraisemblance diminue considérablement après l'ajout des variables explicatives au modèle saturé. Le pseudo R^2 reste, toutefois, faible avec une valeur égale 36%.

Tableau 11 : Tests d'ajustements du modèle

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	37634.8998	5	<.0001
Score	34079.0854	5	<.0001
Wald	19859.2209	5	<.0001

Les tests d'ajustement du modèle montrent que toutes les variables explicatives ont un effet sur la variable dépendante. Ceci est illustré par la statistique de Wald pour chacune des variables. En effet, la p-value correspondante à chaque variable est inférieure au seuil de signification $\alpha = 0.05$.

Tableau 12 : Estimation des coefficients du modèle de capital

Analysis of Maximum Likelihood Estimates								
Parameter		rachat	DF	Estimates	Standard Error	Wald Chi-Square	Pr > ChiSq	Exp(Est)
Intercept		rachat	1	5.1541	0.1273	1638.0651	<.0001	173.143
age_employe		rachat	1	-0.1318	0.00120	11986.7880	<.0001	0.877
Log (anciennete)		rachat	1	-1.7667	0.0151	13769.1801	<.0001	0.171
CRMCCO_c	000.00=<500	rachat	1	2.2454	0.1120	401.5776	<.0001	9.444
CRMCCO_c	001.500=<10000	rachat	1	2.6614	0.1092	593.9753	<.0001	14.316
CRMCCO_c	012.10000=<100000	rachat	1	1.8741	0.1090	295.3637	<.0001	6.515

Les estimations des coefficients du modèle sont toutes significatives au seuil $\alpha = 0.05$. Les coefficients des variables age_employe et log (ancienneté) sont de signes négatifs. Cela signifie que si l'âge à l'embauche ou l'ancienneté augmentent d'une unité, la chance de choisir l'option rachat par rapport à l'individu de référence, diminue respectivement de 13% et 83%. Cependant, le passage de la tranche de cotisation de référence aux autres tranches ne fait qu'augmenter cette chance soit de 9,44 pour la tranche « 000.00=<500 », ou de 14,316 pour la tranche « 001.500=<10000 », ou encore 6,5 pour la tranche « 012.10000=<100000 ».

L'individu de référence est un assuré ayant choisi l'option Non rachat avec un capital dans la tranche supérieure « >=100000 », un âge d'embauche et une ancienneté égale à un an.

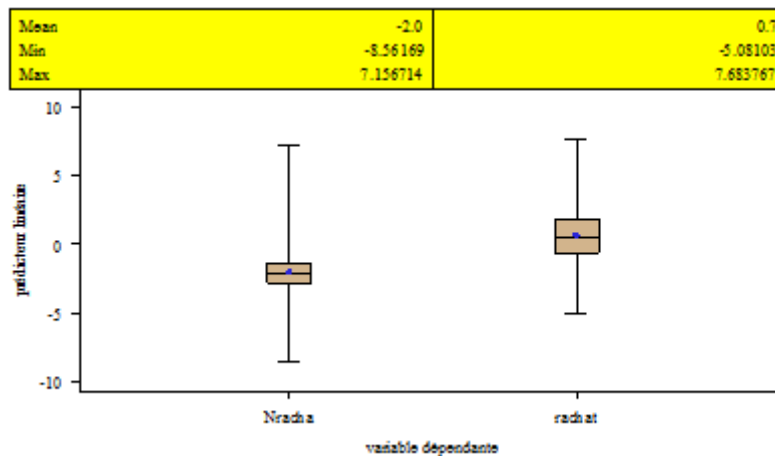
Nous étudions la stabilité du modèle à travers les éléments présentés pour la partie de modélisation de l'option capital.

Tableau 13 : Tableau croisé dressant les valeurs observées contre les valeurs prédites

Table of _FROM_ by _INTO_			
FROM (Formatted Value of the Observed Response)	_INTO_ (Formatted Value of the Predicted Response)		Total
Frequency Percent Row Pct Col Pct	Nracha	rachat	
Nracha	54823 64.99 92.72 84.86	4305 5.10 7.28 21.79	59128 70.09
rachat	9778 11.59 38.75 15.14	15454 18.32 61.25 78.21	25232 29.91
Total	64601 76.58	19759 23.42	84360 100.00

Le tableau ci-dessus, nous indiquent qu'il y a une prédiction du modèle avec un taux de concordance de près de 90%. La statistique C est de l'ordre de $0.89 \approx 1$. 93% des assurés, qui effectivement, n'ayant pas choisi l'option rachat sont correctement classés. Seul 7% sont mal classés par le modèle comme étant des rachats. D'autre part, 39% des assurés qui ont effectivement choisi l'option rachat, ne sont pas classés correctement, en revanche, 61% de ces assurés sont bien classés.

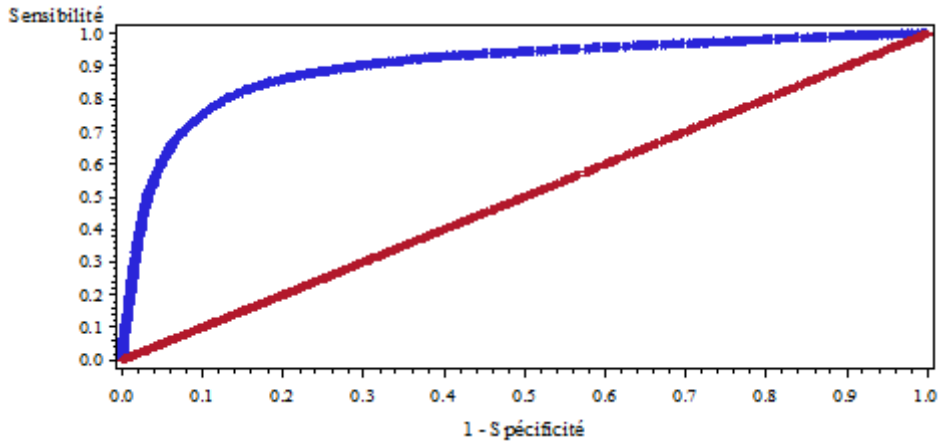
Graphique 9 : Répartition du logit(p) en fonction de l'option rachat



Il est à remarquer que les variables explicatives sont discriminantes. Les deux boîtes à moustaches sont dans des niveaux différents, ainsi que l'allure de la distribution puisque l'écart inter quantile entre les deux groupes est différent.

En effet, celui du groupe NRachat est plus petit que celui du groupe Rachat.

Graphique 10 : La courbe ROC pour l'option rachat



La courbe ROC ci-dessus confirme les résultats précités. Ainsi, la surface entre la courbe ROC (en bleu) et la première bissectrice (Rouge) est assez élevée. Ce qui veut dire que l'ajustement du modèle est acceptable.

Nous récupérons les valeurs prédites de chacun des deux modèles afin de juger si le modèle ajustera bien les autres options qui restent minoritaires. Nous allons croiser les valeurs prédites de chaque modèle pour conclure de l'évolution de la prédiction par rapport aux valeurs observées.

Tableau 14 : Croisement des valeurs observées de l'historique des options rachat et capital

fréquence	capital	Ncapital	Total
Nrachat	577351	400617	977968
rachat	0	456974	456974
Total	577351	857591	1434942

Tableau 15 : Croisement des valeurs prédites sur l'historique des options rachat et capital

fréquence	capital	Ncapital	Total
Nrachat	660257	257043	1068486
rachat	0	357855	357855
Total	660257	648251	1434942

Nous remarquons que les prédictions surestiment la fréquence pour l'option capital. En effet, dans la réalité, on a eu 577351 sorties au titre de l'option capital, alors que le modèle dégage 660257 sorties, soit une surestimation de 14.4%.

Cependant les autres options sont pénalisées par le modèle de prédiction. En effet, les assurés ayant choisi l'option contre-assurance ou l'option rente-viagère sont pénalisés par le modèle de l'ordre de 36% pour l'option rachat, ils sont pénalisés d'un coefficient de l'ordre de 22%.

Nous pouvons ainsi conclure que vu le poids de l'option capital dans notre historique, on tend toujours à trouver des individus qui ont choisi des options hors l'option capital et en même temps trouver des individus avec les mêmes caractéristiques mais qui ont choisi l'option capital. Cela signifie que tous les assurés sont censés sortis en option Capital et toute observation hors celle-là est une erreur (aléa).

2.4.2 Prédictions

Nous déroulons notre modèle estimé sur les contrats en cours (actifs) afin d'estimer le nombre d'assurés dans chacun des groupes capital ou rachat. Nous aurons deux colonnes, une qui nous donne les prédictions en utilisant le modèle qui a comme variable dépendante « Capital » et une seconde pour le modèle qui a comme variable dépendante « Rachat ». Les assurés ayant opté pour l'option rachat seront considérés comme des profils qui effectueront des rachats anticipés tandis que ceux qui ont opté pour l'option capital seront considérés comme des profils qui vont liquider leurs retraites en option capital. Enfin, ceux qui seront classés dans les groupes « Nrachat » et « Ncapital » seront répartis entre l'option contre-assurance et la rente viagère.

Tableau 16 : Croisement des valeurs prédites de l'option de sortie des contrats en cours

Fréquence	Nrachat	rachat	Total
Capital	100869	0	100869
Ncapital	13437	49683	63120
Total	114307	49683	163990

Nous remarquons que 100869 assurés prédits ayant choisi l'option capitale sont tous classés comme des assurés qui n'effectueront pas des rachats. Tandis que pour les assurés prédits dans l'option rachat qui sont de l'ordre de 49683 sont en même temps classés comme étant dans l'option « Ncapital ». Nous retrouvons, cependant, quelques 13437 cas qui figurent simultanément dans les deux groupes « Nrachat » et « Ncapital ». Ceux-là nous allons essayer de les classer dans l'un ou l'autre des groupes « Contre-assurance » et « Rente viagère ».

Nous effectuerons un modèle de régression logistique sur notre historique mais uniquement l'historique composé des options « CAS » et « RVR ». Nous allons estimer notre modèle et analyser sa stabilité afin de procéder à des prédictions sur les 13437 cas restants.

2.4.3 Modélisation de la rente viagère et la contre assurance

2.4.3.1 Le modèle

Variable dépendante qui est le type de sortie qu'on note SICTYP avec :

$$SICTYP : \begin{cases} \text{Rente viagère (RVR) avec une probabilité } p \\ \text{Contre - assurance (CAS) avec une probabilité } 1 - p \end{cases}$$

Le modèle retenue est sous la forme :

$$\text{logit}(p) = \text{age_embauche} + \log(\text{ancienneté}) + \text{CRMCCO_c}$$

Il est le même qui a été proposé précédemment, donc il y a pas lieu d'étudier la corrélation entre les variables explicatives.

Nous effectuons les estimations à l'aide du logiciel SAS afin de trouver les résultats espérés. Les sorties sont dans l'Annexe.

Les estimations du modèle ainsi que le coefficient de concordance entre les valeurs prédites et les valeurs observées sont bonnes. Nous déroulons le modèle estimé sur les 13437 contrats en cours qui existent dans les groupes « Nrachat » et « Ncapital » :

Nous n'obtenons aucune prédiction dans la classe des rentes viagères.

Nous concluons que le modèle que nous avons estimé ne nous donne pas les résultats souhaités malgré les bonnes mesures statistiques d'ajustement qu'il a montrées.

Pour prédire les assurés dans les classes de la rente viagère et celle de la contre-assurance, nous allons nous appuyer sur les statistiques descriptives des 13437 cas identifiés.

2.4.4 Estimation par arbres de décision

Nous avons un échantillon de 6012 individus, avec les variables suivantes « SICTYP », age_embauche, log (anciennete) et CRMCCO_c. La variable SICTYP est une variable avec deux modalités : CAS et RVR soit contre assurance et rente viagère. Nous prenons comme échantillon d'apprentissage 80% de notre base initiale et nous laissons 20% pour le test.

Nous effectuons la construction de l'arbre à l'aide du logiciel SPSS.

Nous construisons notre arbre à l'aide de SPSS en utilisant la méthode CHAID (voir en Annexe l'allure de l'arbre). Nous obtenons ainsi les seuils d'affectation aux classes initiales. L'arbre s'étend sur trois niveaux et ce en introduisant les trois variables explicatives déjà mentionnées.

Pour les assurés ayant un cumul de cotisation ≥ 100000 , la variable CRMCCO_c est la seule variable significative de classement. Pour cette modalité, près de 72% des assurés choisissent l'option rente viagère. Pour les autres modalités, on a besoin d'autres variables explicatives afin de bien classer les assurés.

La seconde variable de classement pour les assurés ayant un cumul de cotisation inférieur à 500 est l'âge à l'embauche. Cependant, l'ajout de cette variable à ce niveau ne permet d'ajouter aucune information au classement déjà effectué par la variable CRMCCO_c.

Pour les autres modalités de la variable CRMCCO_c, notamment les tranches :] 500, 10000] et] 10000,100000], on a besoin de plus d'une variable explicative pour trouver le bon classement. La seconde variable significative est le log (ancienneté). Enfin, pour

certaines feuilles de l'arbre on a encore besoin d'une troisième variable afin de trouver un classement optimal. (Voir Arbre en Annexe).

Sur le tableau de l'arbre « Voir Annexe » nous pouvons remarquer que sur tous les nœuds de l'arbre l'ajout d'une variable indépendante est significative et ce à l'aide de la statistique de χ^2 qui reste égale à 0, inférieure au seuil de 0.0001, (fixé par SPSS), sauf pour les nœuds 20, 21, 22 et 23 ou l'ajout d'une autre variable n'a aucun apport pour distinguer la population au niveau de ces nœuds.

Tableau 17 : Test d'ajustement de l'arbre des échantillons

Risque		
Echantillon	Estimation	Erreur standard
Apprentissage	0,276	0,006
Test	0,279	0,013
Méthode de développement : CHAID Variable dépendante: SICTYP		

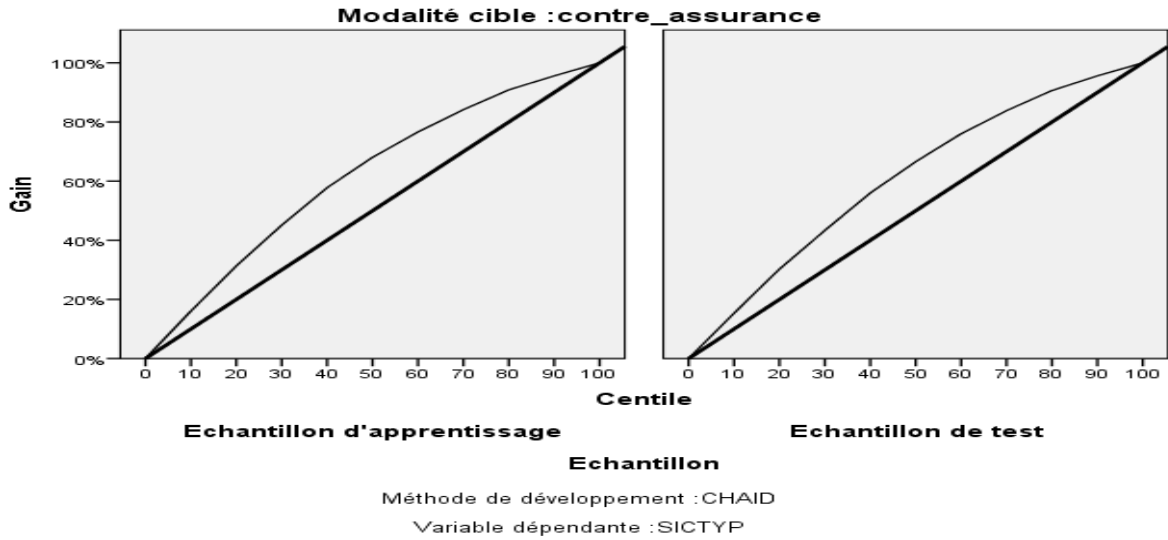
Tableau 18 : Matrice de confusion pour les échantillons d'apprentissage et du test

Echantillon	Observations	Classification		
		contre_assurance	rente_viagere	Pourcentage correct
Apprentissage	contre_assurance	2264	666	77,3%
	rente_viagere	667	1224	64,7%
	Pourcentage global	60,8%	39,2%	72,4%
Test	contre_assurance	563	171	76,7%
	rente_viagere	161	296	64,8%
	Pourcentage global	60,8%	39,2%	72,1%
Méthode de développement : CHAID Variable dépendante : SICTYP				

Les tableaux ci-dessus nous permettent d'évaluer rapidement l'efficacité du model. L'estimation du risque (premier tableaux) qui est de 27% nous renseigne que les modalités prévues par le model sont fausses dans 27% des cas. Cela est vérifié aussi pour l'échantillon du test.

Le second tableau confirme les estimations données par le premier tableau. En effet, le second tableau nous dit que 72% des modalités observées sont correctement classées. Cela est le cas aussi pour l'échantillon de test. Seulement quelques 28% de ces modalités sont mal classés.

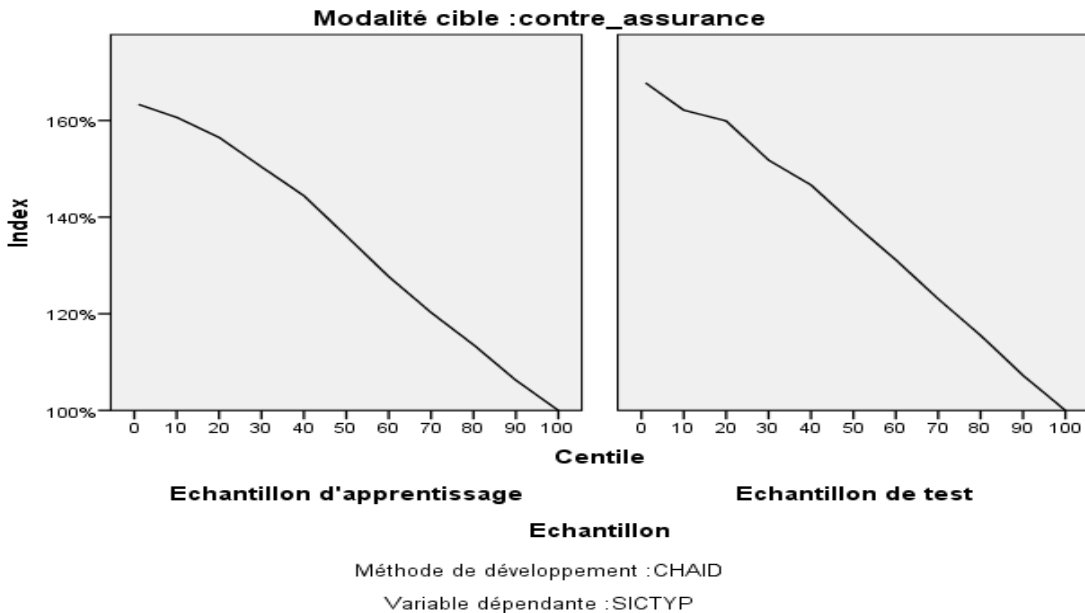
Graphique 11 : La courbe de gain du modèle développé par l'arbre de décision.



Le diagramme de gain indique que le modèle est acceptable. En effet, le diagramme de gain des bons modèles présentent une hausse rapide en direction de la valeur 100%, puis il se stabilise. Les mauvais modèles suivent la diagonale de référence.

Nous notons que c'est le cas, que ce soit pour l'échantillon d'apprentissage ou pour l'échantillon de test.

Graphique 12 : La courbe Index de l'arbre de décision.



Les diagrammes index des bons modèles débutent au-dessus de 100% et restent dans un niveau élevé afin pendant un certain temps avant de diminuer rapidement jusqu'à 100%. Les diagrammes ne fournissant aucune information restent aux alentours de 100% dans l'intégralité du diagramme. (Source IBM SPSS décision trees 19).

L'explication des bons modèles s'applique parfaitement sur les diagrammes au-dessus, que ce soit celui de l'échantillon d'apprentissage ou celui de l'échantillon de test.

Nous avons jusque-là vérifié le bon ajustement de notre modèle, nous allons alors appliquer ces règles pour prédire les personnes figurant dans chacune des modalités. Nous allons utiliser la table des 13437 cas afin d'effectuer cette prédiction.

Tableau 19 : Résultats de la prévision via le modèle pour la rente viagère et contre-assurance

Type de sorties	Fréquence
Contre_assurance	13158
rente_viagere	278
Total	13437

Ci-dessus, les prédictions pour chacune des 13437 cas restants et objet de ce dernier modèle. Nous pouvons noter le faible poids des sorties en rente viagère par rapport au sorties en contre assurance. Toutefois, les gens susceptibles de quitter le régime par voie de décès est très grand.

3. EVALUATION DES ENGAGEMENTS AXA

3.1 HYPOTHESES

Nous nous intéressons dans cette partie au calcul des réserves mathématiques. Ces réserves prendront compte de l'identification de la loi de sortie qui été faite ci-haut. Nous allons nous baser pour ces calculs sur les formules introduites auparavant dans le chapitre « la certification du système ». Toutefois, nous allons présenter brièvement, ici, les formules sur lesquelles nous allons nous baser.

Nous disposons d'une loi de sortie, présentant la possibilité de quitter le portefeuille via quatre options, la première étant l'option capital, la seconde étant l'option rente viagère, la troisième étant l'option rachat et la dernière étant la contre assurance.

Le calcul des réserves pour chacune des options ne se fait pas de la même manière. En effet, chaque option possède une formule qui lui est propre. Toutefois, des hypothèses seront prises en compte notamment pour les âges supérieurs à 60 ans.

L'âge légal de la retraite est 60 ans, cela est une hypothèse valable pour les assurés n'ayant pas atteint l'âge de 60 ans. Toutefois, il existe certains assurés ayant des âges

supérieurs à 60 ans et qui sont encore dans le portefeuille et cela est vrai pour toutes les options.

Nous supposons, donc, que passé l'âge de 60 ans, chaque assuré est susceptible de quitter le portefeuille l'année suivante. Soit un assuré de 61 ans encore dans le portefeuille (contrat en cours), selon l'hypothèse formulé, cet assuré quittera le portefeuille vers la fin de l'année suivante, c'est-à-dire à l'âge de 62 ans.

Le choix de cette hypothèse reste arbitraire. Les chiffres des réserves mathématiques sont à leurs tours des valeurs différentes de ce que nous avons trouvé en réalité. Nous éviterons toute comparaison avec les réserves calculées au sein du système ainsi que toute description du portefeuille ou encore la projection de ces réserves et ce pour des raisons de confidentialité.

Le calcul a été effectué sous Excel à l'aide des formules suivantes :

Avant 60 ans :

Tableau 20 : Formules utilisées pour le calcul des réserves pour les âges inférieures à 60 ans.

Option Capital	$V_x = R_a * (a_{60} + 0.753) * 1.03 * ({}_{60-x}E_x + 0.0015 * {}_{60-x}a_x)$ Le capital constitutif est calculé par la table PF 60-64, l'actualisation est par la table TV 88-90.
Option Rente viagère	$V_x = R_a * (a_{60} + 0.753) * 1.03 * ({}_{60-x}E_x + 0.0015 * {}_{60-x}a_x)$ Le capital et l'actualisation sont calculés par la table TV 88-90.
La contre-assurance	$V_x = CRMCCO * {}_{60-x}A_x$ Avec CRMCCO : cumul des cotisations.
Le rachat	$V_x = CRMCCO$

Après 60 ans :

Tableau 21 : Formules utilisées pour le calcul des réserves pour les âges inférieures à 60 ans.

Option Capital	$V_x = R_a * (a_x + 0.753) * 1.03 * ({}_1E_x + 0.0015 * {}_1a_x)$ <p>Le capital constitutif est calculé par la table PF 60-64, l'actualisation est par la table TV 88-90.</p>
Option Rente viagère	$V_x = R_a * (a_x + 0.753) * 1.03 * ({}_1E_x + 0.0015 * {}_1a_x)$ <p>Le capital et l'actualisation sont calculés par la table TV 88-90.</p>
La contre-assurance	$V_x = CRMCCO * {}_1A_x$ <p>Avec CRMCCO : cumul des cotisations.</p>
Le rachat	$V_x = CRMCCO$

3.2 RESULTATS

Les résultats sont donnés sous la forme du tableau suivant ou nous avons réparti les réserves en fonction des options :

NB : Les chiffres présentés dans le tableau ne reflète pas la vraie valeur du portefeuille.

Tableau 22 : Répartition des réserves selon les types de sorties.

option Sortie	somme reserves
capital	9 785 643 768,51
Contre_assurance	670 213 406,27
rachat	1 415 154 566,35
rente_viagere	23 701 663,53
Total	11 894 713 404,67

Les réserves globales sont de l'ordre de près de 12 Milliards de DH réparties sur les quatre options de sorties. La répartition n'est pas équitable. La rente viagère reste la plus faible comparativement à la contre assurance et le rachat.

CONCLUSION

Au terme de ce chapitre, nous disposons des prédictions que nous avons estimées en se basant, dans un premier temps, sur le modèle logistique qui avait pour variable cible l'option « capital » et dans un second temps, nous avons utilisé le modèle logistique avec l'option « rachat ». Les résultats de ces modèles sont jugés acceptables vu les tests d'ajustement qui ont été utilisés. Toutefois, le croisement des prédictions des deux modèles (capital et rachat) nous laisse une classe qui a eu « non capital » à travers le modèle pour l'option capital et « non rachat » à travers le modèle du rachat.

Pour remédier à cette contrainte nous avons utilisé la méthode des arbres de décisions afin de trouver les caractéristiques qui nous permettront de classer les assurés ayant des contrats en cours que les deux modèles logistiques n'étaient pas capables de prédire. Le modèle de l'arbre de décision est jugé acceptable. Nous utilisons les règles de classement trouvées par ce modèle afin de faire les prévisions sur les contrats en cours.

Conclusion générale et perspectives

S'il y a quelques de choses de plus important pour un assureur vie, c'est bien évidemment sa solvabilité qui doit veiller à sa garantie moyennant un bon calcul des réserves mathématiques pour assurer un suivi régulier de l'équilibre de son portefeuille. C'est cela qui permet à l'assureur d'arrêter ses engagements futurs envers ces assurés et de dresser sa stratégie d'investissement.

Dans les contrats des garanties vie ou de retraite où les modalités de liquidation sont multiples, le calcul des réserves est toujours taché de doute. Ce dernier provient, essentiellement, du comportement du portefeuille de l'assureur : est ce que les assurés ont le même comportement lors de la liquidation de leurs retraites ?

C'est dans ce contexte de prédiction du comportement des assurés et de calcul des réserves mathématiques que s'inscrit notre étude, dans le but d'approcher ce comportement futur et de prédire quel serait le comportement d'un assuré donné.

En effet, nous avons utilisé, à cette fin, un premier modèle de la régression logistique afin de prédire les sorties qui s'effectueront en option capital. Dans ce modèle nous essayons d'expliquer la variable binaire : capital qui prend deux valeurs, capital et non-capital par deux variables continues et une autre catégorielle. Nous avons étudié la corrélation entre les variables explicatives de ce modèle et nous avons conclu qu'il n'existe pas de liaisons entre ces variables. Par la suite, nous avons utilisé le même modèle pour prédire les sorties en option rachat. Les résultats des estimations des coefficients des deux modèles étaient bons ainsi que les tests d'ajustement qui traduisaient des taux de concordance de l'ordre de 80%.

Nous avons croisé les prédictions des rachats et des sorties en capital à travers l'application des modèles estimés afin de vérifier la nullité de la fréquence dans la classe (Rachat, Capital), qui traduit la viabilité du modèle utilisé.

Cependant, il existe des non classements des assurés dont nous ne connaissons pas le type de sortie par aucun des deux modèles (classés comme des non-rachats par le modèle de rachat et comme non-capital par le modèle du capital). Ceux-là sont alors nécessairement dans la catégorie des assurés qui feront des sorties dans l'une des deux options : rente viagère ou contre assurance.

Pour cela nous avons effectué une approche descriptive via les arbres de décision (voir

Annexe 7, page 80) afin de trouver les règles selon lesquelles nous pouvons classer ses assurés. Les résultats de l'arbre de décision étaient bons et ont présenté un taux de concordance de près de 72%. Nous avons utilisé les règles données par l'arbre de décisions afin de prédire les types de sortie des assurés que nous avons identifiés comme non-classement par les deux modèles logistiques.

Enfin, nous avons utilisé les prédictions déduites par les deux modèles de la régression logistique et le modèle de l'arbre de décision pour calculer les provisions mathématiques qu'AXA devra constituer. En effet, nous avons classé les résultats des prédictions en quatre groupes définis auparavant, un groupe des assurés prédits pour une liquidation en capital, un groupe des assurés prédits pour une liquidation en rachat et deux autres pour des liquidation en rente viagère et en contre-assurance. Le calcul de la provision a une démarche spécifique pour chacun de ces groupes.

Ce travail de recherche reste perfectible. Il peut être complété et approfondi davantage. Les principales perspectives de recherche qu'on peut évoquer à titre de suggestion sont :

- La modélisation de la durée de service pour les contrats en cours à l'aide du modèle du Kaplan Meier.
- La projection des provisions mathématiques jusqu'au dernier assuré vivant.

Références bibliographiques

❖ Bibliographie

- [1] David W. Hosmer, JR ; Stanley Lemeshow ; Rodney X. Sturdivant; Applied Logistic Regression; Third Edition; 2013; wiley.
- [2] Paul D. Allison ; Logistic Regression Using The SAS System : Theory and Application; 1999,SAS Institute and Wiley
- [3] Pierre Petauton ; Théorie et pratique de l'assurance vie ; 3eme édition ; 2004 ; Dunod.
- [4] R. Arboretti Giancristofaro, L. Salmaso; Model Performance Analysis And Model Validation In Logistic Regression; Statistica, anno LXIII, n. 2, 2003
- [5] Sentence arbitrale (CIMR-FMSAR) ; 1992.
- [6] Statuts Règlement Intérieur, 2 Décembre 2008.
- [7] Stéphane Tufféry, Data mining et statistique décisionnelle, 2005, Technip.
- [8] Thierry Poincelin, Calculs Elémentaires de l'Assurance Vie ; 2003 ; Economica.

❖ Webographie

- [9] WWW.cimr.ma
- [10] Document: CIMR: La sentence arbitrale de M Jouahri ; Edition Num. 21 du 19/03/1992 :adresse <http://www.leconomiste.com/article/document-cimr-la-sentence-arbitrale-de-m-jouahri>
- [11] www.support.sas.com

Annexes

Annexe 1 : Table des Régimes de retraite de base au Maroc.....	69
Annexe 2 : Table des Régimes de retraite complémentaires au Maroc.....	70
Annexe 3 : Table des coefficients d'anticipation et de prorogation de la CIMR.....	71
Annexe 4 : Table de mortalité PF 60-64 avec commutations	71
Annexe 5 : Table de mortalité TV 88-90 avec commutations.....	74
Annexe 6 : Sorties SAS pour le modèle de la rente viagère et contre assurance	77
Annexe 7 : Arbre de décision pour le modèle de la rente viagère et de la contre- assurance.....	80

Annexe 1 : Table des Régimes de retraite de base au Maroc

Organisme	CMR	CNSS	RCAR
Création	1930	1959	1977
Population	<ul style="list-style-type: none"> fonctionnaires titulaires et stagiaires de l'Etat et des collectivités locales et de certains établissements publics Personnel des FAR, forces auxiliaires et de la Gendarmerie Royale Anciens résistants et leurs veuves, descendantes et ascendantes ; Les fonctionnaires du parlement. 	<p>Les salariés des :</p> <ul style="list-style-type: none"> entreprises privées de l'industrie, du commerce et des services entreprises artisanales associations et syndicats Marins pêcheurs travailleurs salariés de l'agriculture (hors AF). travailleurs occasionnels ou temporaires du secteur privé. 	<ul style="list-style-type: none"> le personnel des organismes soumis au contrôle financier de l'Etat. le personnel contractuel de droit commun, temporaire, journalier et occasionnel de l'Etat et des collectivités locales.
Mode de fonctionnement	Répartition, Système de la prime échelonnée (10 ans min).	Répartition (Système de la prime échelonnée (5 ans min)).	système mixte : 1/3 répartition 2/3 capitalisation
Prestations garanties	<ul style="list-style-type: none"> pensions de vieillesse pensions d'invalidité pensions de survivants (pension de veuve (veuf) et d'orphelins) remboursement des cotisations. allocations familiales pour les retraités 	<ul style="list-style-type: none"> les prestations familiales (allocations familiales, aide sanitaire familiale), les prestations dites à court terme (indemnités journalières en cas de maladie ou de maternité, allocation au décès, remboursement du congé de naissance) les prestations dites à long terme (pension d'invalidité, de vieillesse, de survivants). 	<ul style="list-style-type: none"> des pensions de vieillesse des pensions d'invalidité des pensions d'ayants causent des allocations familiales au titre des enfants à charge pour les bénéficiaires de pensions.
Cotisation	<p>Régime Civil :</p> <ul style="list-style-type: none"> Taux de cotisation est de 20% supporté à parts égales entre l'employeur et l'affilié (14% en 2003). <p>Régime militaire :</p> <ul style="list-style-type: none"> Part Salariale est de 7%, tandis que patronale est de 14%. 	<ul style="list-style-type: none"> 11,89% de la rémunération brute plafonnée à 6.000 dh/mois ; à charge des employeurs et des salariés respectivement à hauteur de 2/3 et 1/3, calculées sur la base d'un taux. 	Les cotisations salariales et patronales sont respectivement de 6% et 12 % de la rémunération plafonnée (12 900 dirhams/mois en 2008).
Jouissance de la retraite	<ul style="list-style-type: none"> 60 ans pour l'ensemble des affiliés. 62 ans pour les fonctionnaires de la sécurité ; 65 ans pour professeurs chercheurs de l'enseignement supérieur ; 66 ans pour les magistrats <p><u>Mise en retraite sur demande :</u> Durée minimale de service :</p> <ul style="list-style-type: none"> Femme : 15 ans de service Homme : 21 ans de service <p>Avec Autorisation</p>	<ul style="list-style-type: none"> 60 ans pour une période d'assurance (cotisation) d'au moins 3240 jours. 55 ans pour les mineurs qui justifient avoir travaillé au fond pendant au moins 5 ans. Possibilité d'anticipation entre 55 ans et 60 ans à condition de disposer d'au moins 3240 jrs de cotisation et avec accord de l'employeur. 	<ul style="list-style-type: none"> 60 ans pour retraite normale anticipée dès l'âge de 55 ans ou après 21 ans de services valables et sans condition d'âge
Plan de calcul de	2,5% *N*assiette liquidation	Pour 3240 jours d'assurance est	2%*N*Salaire moyen de

Pension	N : nombre d'années de cotisations valides et validées ; Assiette de liquidation : Dernier salaire perçu.	égal à 50% du salaire des quatre-vingt-seize derniers mois d'assurance + 1% du meme montant pour chaque période d'assurance de 216 jours accomplis en sus des 3240, sans toutefois dépasser 70% du salaire plafonné.	carrière revalorisée (sans dépasser le plafond).
----------------	---	--	--

Annexe 2 : Table des Régimes de retraite complémentaires au Maroc

Organisme	CIMR normal	CIMR complémentaire	CMR ATTAKMILI	Assurances privées	RCAR Complémentaire
Création	1949	1993	2005	-	1977
Population	salariés du secteur privé	Les employeurs adhérant à la CIMR au niveau maximum de contribution (6%)	Affiliés du régime de base CMR	Particuliers, groupes de personnes.	Assurés RCAR dont le salaire est supérieur au plafond.
Mode de fonctionnement	<u>Avant 2003 :</u> <ul style="list-style-type: none"> Part patronale : répartition Part salariale gérée par les compagnies d'assurance : en capitalisation <u>Après 2003 :</u> <ul style="list-style-type: none"> cotisations salariales gérées en répartition transformées en points 	octroi annuel des points, la valeur du point à la liquidation et les conditions de liquidation sont les mêmes que dans le cas de la retraite normale.	Capitalisation individuelle	Capitalisation	<ul style="list-style-type: none"> 1/2 capitalisation pour part salariale 1/2 répartition pour part patronale Régime en points
Prestations garanties	<ul style="list-style-type: none"> pensions de retraite pension d'invalidité (sous forme de points gratuits) et d'ayant cause 	Pension normales majorées de 16,67 %, 25% ou de 33,33%.	<ul style="list-style-type: none"> Capital retraite Rente certaine Rente viagère Mixte Si Affiliation entre 5 ans et 10 ans alors récupérer 97% des droits acquis Si >10 ans alors 100% des droits acquis 	<ul style="list-style-type: none"> capital rente 	<ul style="list-style-type: none"> Pension vieillesse Pension invalidité Pension décès Pension viagère
Cotisation	<ul style="list-style-type: none"> Les cotisations calculées sur la base d'une assiette plafonnée (variable d'année en année), sont transformées en 	Le taux de cotisation sera, pour chacune des deux parties patronale et salariale majoré de 1%, 1,5%, 2%, c'est-à-dire un taux de :	Au choix des Assurés	Définies par la réglementation interne de chaque compagnie	<ul style="list-style-type: none"> Cotisations salariales et contributions patronales dont rémunération est supérieure au plafond du régime base. Les taux sont fixés à 3% pour chacune

	points	7% ; 7,5% ou 8% du salaire où le plafond égal à 1,5 fois celui du régime.			des deux parts.
Jouissance de retraite	<ul style="list-style-type: none"> • 60 ans retraite normale • anticiper à 50 ans • Prorogation à 65 ans. 	Mêmes conditions que la CIMR normale.	<ul style="list-style-type: none"> • Similaire à celle du régime de base sauf : Anticipation à 5 ans (max) avant retraite légale • Prorogation • Possibilité de rachat si 5 ans d'affiliation minimum est vérifié. 	Définies par la réglementation interne de la compagnie	<ul style="list-style-type: none"> • Total points >1000 points • 60 ans Age normal • 55 ans si anticipation • Cessation de services pour incapacité totale • Décès de l'assuré

Annexe 3 : Table des coefficients d'anticipation et de prorogation de la CIMR

nombre d'années	Coefficients de prorogation de la CIMR	Coefficients d'anticipation de la CIMR
1	1,05	0,93
2	1,10	0,87
3	1,15	0,81
4	1,2	0,76
5	1,25	0,72
6	1,25	0,66
7	1,25	0,61
8	1,25	0,54
9	1,25	0,47
10	1,25	0,4

Annexe 4 : Table de mortalité PF 60-64 avec commutations

taux technique /1993	0,045				
x	PF 60-64	Dx	Nx	Cx	Mx
0	1000000	1000000	21721148,3	18087,5092	66077,9312
1	981510	939244,019	20721148,3	1862,85324	47990,4219
2	979520	896975,802	19781904,3	814,278855	46127,5687
3	978611	857553,496	18884928,5	522,905029	45313,2898
4	978001	820113,833	18027375	393,74761	44790,3848
5	977521	784412,749	17207261,2	313,993309	44396,6372
6	977121	750327,052	16422848,4	255,401256	44082,6439
7	976781	717766,476	15672521,4	215,649808	43827,2426
8	976481	686646,916	14954754,9	186,414986	43611,5928

9	976210	656896,031	14268108	163,9059	43425,1778
10	975961	628448,304	13611212	151,808466	43261,2719
11	975720	601237,433	12982763,7	144,668474	43109,4635
12	975480	575205,307	12381526,2	143,630184	42964,795
13	975231	550295,197	11806320,9	149,036912	42821,1648
14	974961	526452,482	11256025,7	163,747803	42672,1279
15	974651	503622,096	10729573,2	181,970085	42508,3801
16	974291	481757,011	10225951,1	198,319338	42326,41
17	973881	460817,491	9744194,14	218,014723	42128,0907
18	973410	440760,407	9283376,64	230,330776	41910,0759
19	972890	421554,977	8842616,24	241,605709	41679,7452
20	972320	403165,545	8421061,26	243,370142	41438,1395
21	971720	385566,277	8017895,72	240,264941	41194,7693
22	971101	368727,91	7632329,44	241,433106	40954,5044
23	970451	352613,497	7263601,53	242,055127	40713,0713
24	969770	337192,398	6910988,03	244,216683	40471,0161
25	969052	322433,25	6573795,63	246,394196	40226,7995
26	968295	308307,534	6251362,38	248,865722	39980,4053
27	967496	294787,685	5943054,85	251,263605	39731,5395
28	966653	281847,684	5648267,17	254,419607	39480,2759
29	965761	269461,821	5366419,48	256,837867	39225,8563
30	964820	257606,955	5096957,66	259,882013	38969,0185
31	963825	246259,608	4839350,71	259,688309	38709,1364
32	962786	235401,091	4593091,1	260,225263	38449,4481
33	961698	225009,642	4357690,01	261,607687	38189,2229
34	960555	215064,318	4132680,36	263,921655	37927,6152
35	959350	205544,999	3917616,05	266,389584	37663,6935
36	958079	196433,188	3712071,05	269,960646	37397,3039
37	956733	187710,258	3515637,86	274,457527	37127,3433
38	955303	179358,558	3327927,6	279,168495	36852,8858
39	953783	171361,893	3148569,04	285,425357	36573,7173
40	952159	163703,46	2977207,15	291,802975	36288,2919
41	950424	156368,578	2813503,69	299,516202	35996,4889
42	948563	149342,006	2657135,11	307,872182	35696,9727
43	946564	142609,84	2507793,11	317,163815	35389,1006
44	944412	136158,487	2365183,27	327,340857	35071,9367
45	942091	129974,988	2229024,78	338,617537	34744,5959
46	939582	124046,732	2099049,79	350,640703	34405,9784
47	936867	118361,999	1975003,06	363,842988	34055,3377
48	933923	112909,148	1856641,06	377,97814	33691,4947
49	930727	107677,281	1743731,91	393,163721	33313,5165
50	927253	102655,855	1636054,63	409,48124	32920,3528
51	923472	97834,7004	1533398,78	426,980713	32510,8716
52	919352	93204,0366	1435564,08	445,585623	32083,8909
53	914859	88754,5796	1342360,04	465,307823	31638,3052
54	909956	84477,433	1253605,46	486,137825	31172,9974

55	904603	80364,0933	1169128,03	508,134841	30686,8596
56	898756	76406,3648	1088763,93	531,327725	30178,7247
57	892367	72596,3768	1012357,57	555,719123	29647,397
58	885384	68926,5948	939761,191	581,213068	29091,6779
59	877752	65389,9027	870834,596	607,78051	28510,4648
60	869412	61979,5193	805444,694	635,305778	27902,6843
61	860302	58689,0683	743465,174	663,938065	27267,3785
62	850353	55512,3024	684776,106	693,268842	26603,4405
63	839497	52443,6428	629263,804	723,425632	25910,1716
64	827659	49477,626	576820,161	754,144018	25186,746
65	814763	46609,2827	527342,535	785,184288	24432,602
66	800732	43834,092	480733,252	816,383303	23647,4177
67	785487	41147,8873	436899,16	847,487658	22831,0344
68	768949	38546,9292	395751,273	878,077168	21983,5467
69	751043	36028,0504	357204,344	907,886264	21105,4696
70	731696	33588,4799	321176,293	936,418676	20197,5833
71	710843	31226,0513	287587,813	963,173667	19261,1646
72	688429	28939,1816	256361,762	987,656174	18297,991
73	664411	26726,8399	227422,58	1009,22784	17310,3348
74	638764	24588,6644	200695,74	1027,29848	16301,1069
75	611483	22524,8861	176107,076	1041,07651	15273,8085
76	582592	20536,5006	153582,19	1049,96982	14232,7319
77	552143	18625,0397	133045,689	1053,13086	13182,7621
78	520228	16792,7982	114420,65	1049,96763	12129,6313
79	486977	15042,5514	97627,8515	1039,68485	11079,6636
80	452570	13377,7331	82585,3001	1021,86347	10039,9788
81	417231	11802,0389	69207,567	995,92884	9018,11531
82	381239	10319,5677	57405,5281	961,674195	8022,18647
83	344921	8934,44397	47085,9604	918,995436	7060,51228
84	308653	7650,71642	38151,5164	868,19471	6141,51684
85	272848	6471,96391	30500,8	809,739379	5273,32213
86	237951	5401,15395	24028,8361	744,58329	4463,58275
87	204418	4440,19319	18627,6822	673,932946	3718,99946
88	172701	3589,72594	14187,489	599,385607	3045,06652
89	143223	2848,80627	10597,763	522,809545	2445,68091
90	116354	2214,70139	7748,95677	446,224229	1922,87136
91	92389	1682,82066	5534,25538	371,862062	1476,64713
92	71519	1246,58734	3851,43472	301,712776	1104,78507
93	53824	897,761461	2604,84738	237,66605	803,072297
94	39258	626,609347	1707,08592	181,183345	565,406246
95	27654	422,386828	1080,47657	133,143726	384,222902
96	18743	273,952471	658,089745	93,9813338	251,079176
97	12170	170,219929	384,137274	63,4315186	157,097842
98	7534	100,839144	213,917345	40,7198135	93,6663234
99	4424	56,6633653	113,078201	24,7329688	52,94651
100	2450	30,0287352	56,4148355	14,1360026	28,2135412

101	1271	14,907342	26,3861003	7,54957342	14,0775387
102	613	6,88016493	11,4787583	3,73300994	6,52796523
103	273	2,93214019	4,59859339	1,70207602	2,79495529
104	111	1,14085059	1,6664532	0,70379421	1,09287927
105	41	0,40324906	0,5256026	0,26939491	0,38908506
106	13	0,12235355	0,12235355	0,11969015	0,11969015
107	0	0	0	0	0

Annexe 5 : Table de mortalité TV 88-90 avec commutations

taux technique /1993	0,045				
x	TV 88-90	Dx	Nx	Cx	Mx
0	1000000	1000000	22239878,1	6338,94321	43244,4524
1	993520	950736,842	21239878,1	542,942149	36905,5092
2	992940	909264,898	20289141,3	295,612786	36362,567
3	992610	869820,772	19379876,4	214,30534	36066,9542
4	992360	832154,735	18510055,6	180,467654	35852,6489
5	992140	796143,781	17677900,9	156,996655	35672,1812
6	991940	761706,499	16881757,1	127,700628	35515,1846
7	991770	728780,819	16120050,6	115,013231	35387,484
8	991610	697285,404	15391269,8	110,060508	35272,4707
9	991450	667151,095	14693984,4	105,32106	35162,4102
10	991290	638319,072	14026833,3	107,08481	35057,0892
11	991120	610726,894	13388514,2	96,4456496	34950,0043
12	990960	584333,304	12777787,3	86,5242072	34853,5587
13	990810	559085,985	12193454	104,877827	34767,0345
14	990620	534907,917	11634368	110,925931	34662,1567
15	990410	511765,093	11099460,1	116,258666	34551,2307
16	990180	489613,634	10587695	140,274654	34434,9721
17	989890	468392,572	10098081,4	157,377932	34294,6974
18	989550	448068,605	9629688,8	186,036396	34137,3195
19	989130	428591,798	9181620,19	186,502652	33951,2831
20	988690	409953,249	8753028,39	186,583776	33764,7804
21	988230	392117,238	8343075,14	174,667566	33578,1967
22	987780	375060,941	7950957,91	163,431641	33403,5291
23	987340	358750,117	7575896,97	159,948322	33240,0975
24	986890	343145,081	7217146,85	166,665981	33080,1491
25	986400	328205,461	6874001,77	162,743855	32913,4831
26	985900	313913,01	6545796,31	165,079891	32750,7393
27	985370	300233,738	6231883,3	163,932364	32585,6594
28	984820	287144,649	5931649,56	154,020838	32421,727
29	984280	274628,9	5644504,91	155,576604	32267,7062
30	983710	262650,585	5369876,01	159,324651	32112,1296

31	983100	251184,417	5107225,43	157,462594	31952,8049
32	982470	240213,828	4856041,01	155,46546	31795,3424
33	981820	229717,61	4615827,18	162,503462	31639,8769
34	981110	219666,498	4386109,57	175,217696	31477,3734
35	980310	210035,772	4166443,07	186,535586	31302,1557
36	979420	200808,694	3956407,3	182,514255	31115,6201
37	978510	191982,889	3755598,61	188,089774	30933,1059
38	977530	183531,687	3563615,72	192,846658	30745,0161
39	976480	175439,76	3380084,03	200,360164	30552,1695
40	975340	167688,939	3204644,27	203,505245	30351,8093
41	974130	160268,809	3036955,33	210,83623	30148,3041
42	972820	153161,035	2876686,52	221,778861	29937,4678
43	971380	146348,632	2723525,49	231,388099	29715,689
44	969810	139820,187	2577176,85	241,168835	29484,3009
45	968100	133563,303	2437356,67	253,72697	29243,132
46	966220	127563,569	2303793,36	255,715872	28989,4051
47	964240	121820,252	2176229,8	254,591221	28733,6892
48	962180	116325,354	2054409,54	263,733183	28479,098
49	959950	111058,136	1938084,19	275,010893	28215,3648
50	957520	106006,704	1827026,05	285,91126	27940,3539
51	954880	101162,135	1721019,35	296,399233	27654,4426
52	952020	96515,9231	1619857,21	307,437221	27358,0434
53	948920	92058,9901	1523341,29	315,076886	27050,6062
54	945600	87786,5091	1431282,3	313,315057	26735,5293
55	942150	83699,734	1343495,79	318,94217	26422,2142
56	938480	79783,4398	1259796,06	333,483202	26103,2721
57	934470	76021,5654	1180012,62	344,58883	25769,7889
58	930140	72410,8216	1103991,05	357,165787	25425,2
59	925450	68943,2613	1031580,23	360,733036	25068,0342
60	920500	65621,5321	962636,97	367,514978	24707,3012
61	915230	62436,2096	897015,438	379,717317	24339,7862
62	909540	59376,1174	834579,228	390,187235	23960,0689
63	903430	56437,5575	775203,111	400,884621	23569,8817
64	896870	53615,0739	718765,554	414,615469	23168,9971
65	889780	50900,7007	665150,48	420,824306	22754,3816
66	882260	48297,1407	614249,779	437,510763	22333,5573
67	874090	45789,3725	565952,638	459,154034	21896,0465
68	865130	43368,4222	520163,266	485,968096	21436,8925
69	855220	41025,4929	476794,844	507,744321	20950,9244
70	844400	38762,1532	435769,351	533,928838	20443,1801
71	832510	36570,6633	397007,198	565,081365	19909,2512
72	819360	34443,0694	360436,534	597,084172	19344,1699
73	804840	32375,7882	325993,465	631,185504	18747,0857
74	788800	30364,1696	293617,677	668,773911	18115,9002
75	771040	28402,4056	263253,507	709,161525	17447,1263
76	751360	26485,611	234851,102	743,106492	16737,9647

77	729810	24618,1518	208365,49	786,672088	15994,8583
78	705970	22788,4922	183747,339	832,054588	15208,1862
79	679620	20993,2272	160958,846	882,04147	14376,1316
80	650430	19226,3714	139965,619	922,710416	13494,0901
81	618520	17495,8167	120739,248	961,008241	12571,3797
82	583790	15802,3194	103243,431	996,944585	11610,3715
83	546140	14146,5936	87441,1117	1010,77335	10613,4269
84	506250	12548,6394	73294,5181	1011,13586	9602,65351
85	464550	11019,1419	60745,8787	1003,55985	8591,51766
86	421300	9562,91908	49726,7368	975,221366	7587,95781
87	377380	8197,12601	40163,8177	934,501087	6612,73644
88	333400	6929,98088	31966,6917	886,532752	5678,23535
89	289800	5764,32596	25036,7108	825,202011	4791,7026
90	247390	4708,86241	19272,3849	751,310146	3966,50059
91	207040	3771,13282	14563,5225	667,284821	3215,19044
92	169590	2955,98019	10792,3896	576,144374	2547,90562
93	135800	2265,08633	7836,40946	480,357581	1971,76125
94	106360	1697,64558	5571,32313	393,157241	1491,40367
95	81180	1239,94224	3873,67755	307,94436	1098,24643
96	60570	885,306577	2633,73532	240,064901	790,302067
97	43780	612,344165	1748,42874	175,408125	550,237166
98	30960	414,385441	1136,08457	119,409871	374,829041
99	21840	279,730537	721,699133	88,3320313	255,41917
100	14790	181,275507	441,968595	62,1072886	167,087139
101	9610	112,714049	260,693088	41,5341273	104,97985
102	5990	67,2303228	147,979038	26,4604528	63,4457226
103	3580	38,4507761	80,7487156	16,0751624	36,9852698
104	2050	21,0697632	42,2979395	9,24986673	20,9101074
105	1130	11,1139374	21,2281763	5,1954733	11,6602407
106	590	5,55296868	10,1142389	2,67001109	6,46476736
107	300	2,70195541	4,56127018	2,64313918	3,79475627
108	140	1,20661485	1,85931478	0,67448528	1,1516171
109	60	0,4948523	0,65269992	0,32272023	0,47713182
110	20	0,15784762	0,15784762	0,15441159	0,15441159
111	0	0	0	0	0

Annexe 6 : Sorties SAS pour le modèle de la rente viagère et contre assurance

Statistiques d'ajustement du modèle		
Critère	Coordonnée à l'origine uniquement	Coordonnée à l'origine et covariables
AIC	8045.989	6917.793
SC	8052.691	6958.002
-2 Log L	8043.989	6905.793
R carré	0.1725	
R carré remis à l'échelle max.	0.2338	

Test de l'hypothèse nulle globale : BETA=0			
Test	Khi-2	DD L	Pr > Khi-2
Rapport de vrais	1138.19 68	5	<.0001
Score	993.173 0	5	<.0001
Wald	806.589 6	5	<.0001

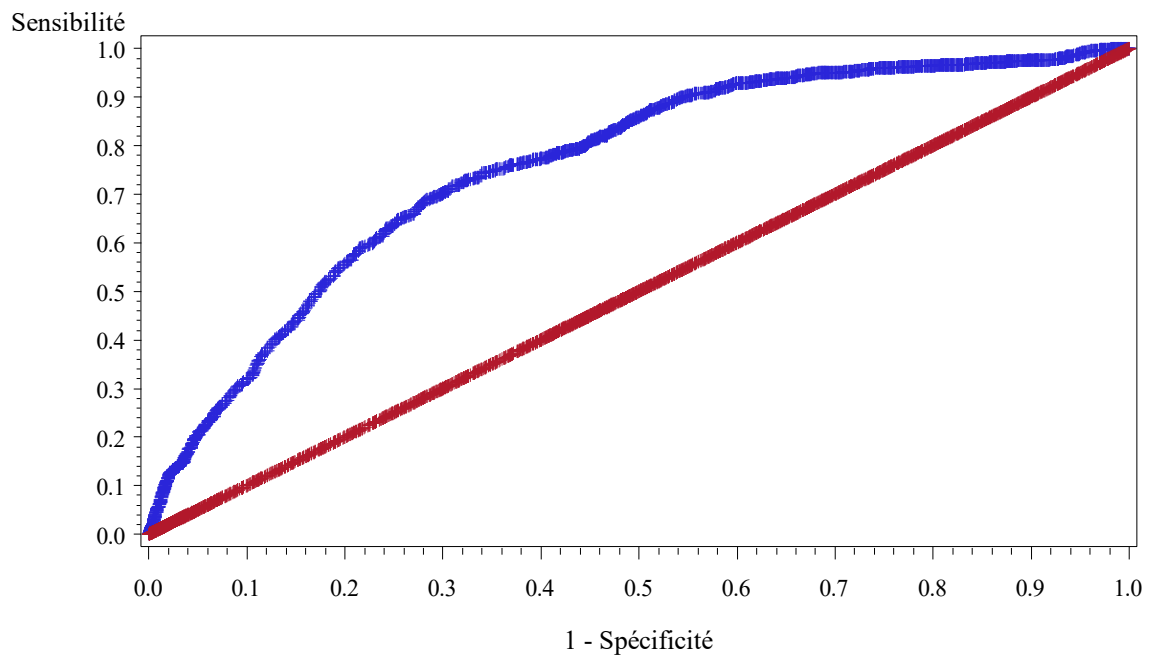
Analyse des effets Type 3			
Effet	DD L	Khi-2 de Wald	Pr > Khi-2
age_embauche	1	110.8183	<.0001
lanciennete	1	217.1319	<.0001
CRMCCO_c	3	476.1411	<.0001

Analyse des estimations de la vraisemblance maximum								
Paramètre		SICTY P	DDL	Valeur estimée	Erreur type	Khi-2 de Wald	Pr > Khi-2	Exp(Est)
Intercept		RVR	1	-3.0600	0.3049	100.7382	<.0001	0.047
age_embauche		RVR	1	0.0397	0.00377	110.8183	<.0001	1.040
lanciennete		RVR	1	0.9441	0.0641	217.1319	<.0001	2.571
CRMCCO_c	000.00=<500	RVR	1	-2.3542	0.1691	193.9059	<.0001	0.095
CRMCCO_c	001.500=<10000	RVR	1	-2.0785	0.1209	295.6710	<.0001	0.125
CRMCCO_c	012.10000=<100000	RVR	1	-0.8890	0.1098	65.5262	<.0001	0.411

Estimations des rapports de cotes					
Effet		SICTY P	Valeur estimée du point	Intervalle de confiance de Wald à 95 %	
age_embauche		RVR	1.040	1.033	1.048
lanciennete		RVR	2.571	2.267	2.915
CRMCCO_c	000.00=<500 vs 026.=>200000	RVR	0.095	0.068	0.132
CRMCCO_c	001.500=<10000 vs 026.=>200000	RVR	0.125	0.099	0.159
CRMCCO_c	012.10000=<100000 vs 026.=>200000	RVR	0.411	0.331	0.510

Association des probabilités prédites et des réponses observées			
Pourcentage concordant	75.6	D de Somers	0.516
Pourcentage discordant	24.0	Gamma	0.518
Pourcentage lié	0.3	Tau-a	0.246
Paires	8603072	c	0.758

Courbe ROC :



Annexe 7 : Arbre de décision pour le modèle de la rente viagère et de la contre-assurance

