

ROYAUME DU MAROC
*_*_*_*_*
HAUT COMMISSARIAT AU PLAN
*_*_*_*_*_*_*_*_*

INSTITUT NATIONAL
DE STATISTIQUE ET D'ECONOMIE APPLIQUEE

INSEA



Projet de Fin d'Etudes

Estimation des provisions pour sinistres à payer pour la branche accident de travail

Préparé par : *Mlle Ghizlane Khouzaimi*
Mlle Amal Hannouda

Sous la direction de : *M. Fouad Marri (INSEA)*
M. Mohammed Boutellaka (MAMDA-MCMA)

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. Fouad Marri (INSEA)*
- *M. Effina Driss (INSEA)*
- *M. Mohammed Boutellaka (MAMDA-MCMA)*

Juin 2012

Dédicaces 1

En préambule de ce travail, je tiens à gratifier mes parents pour l'amour et l'affection qu'ils m'ont prodigués. Ils ont toujours veillé à ce que je ne manque de rien. Mon épanouissement et ma réussite professionnels ne sont tributaires que de leur persévérance dans mon éducation. Pour toutes ces raisons et bien d'autres, je leur dis que je les aime à la folie.

Je tiens absolument à remercier ma sœur Sofia et mon frère Mehdi pour leur soutien moral et pour leur présence aux moments difficiles. Leurs Vannes sont venues illuminées mes journées les plus stressantes. Chers frère et sœur, vous êtes les meilleurs.

Il va de soit de vouloir exprimer sa gratitude envers ses meilleures amies. Ainsi, je désigne Maha Mzari et Fatima Zahra Mahfoudi pour leur dire qu'elles comptent énormément et que je suis extrêmement reconnaissante envers elles. Merci d'avoir été la perle des amies.

Un spécial remerciement à mon binôme Amal pour son sérieux et pour son grand esprit d'équipe qui rendait agréable toute tâche difficile. Ce fût un plaisir de collaborer avec toi.

Ghizlane

Dédicaces 2

Je dédie ce travail

A mes très chers parents

Tous les mots du monde ne sauraient exprimer l'immense amour que je vous porte, ni la profonde gratitude que je vous témoigne pour tous les efforts et les sacrifices que vous n'avez jamais cessé de consentir pour mon instruction et mon bien-être. C'est à travers vos encouragements et vos critiques que je me suis réalisée.

A la mémoire de ma tante Fatima Nekker

Celle qui a su être bonne, gentille et compréhensive avec moi ; celle dont je regrette l'absence à cette étape importante de ma vie ; celle qui me manque terriblement aujourd'hui, ma très chère et adorée tante Fatima.

A mes chers frères Houssam et Nazih

Pour leurs encouragements et leur soutien permanents.

A ma tante Zahra Nekker et Mr Abdallah Rhalimi

J'ai toujours senti que vous êtes ma deuxième famille que j'aime et je respecte. Je vous remercie pour tout ce que vous m'avez apporté.

A mes chères cousines

Celles qui m'ont toujours aidée, écoutée, soutenue et encouragée tout au long de mon parcours, celles qui ont toujours été présentes pour moi, mes très chères cousines Mouna, Loubna, Sahar et Soumia.

À mes Grand Parents

Quoique je puisse dire, je ne peux exprimer mes sentiments d'amour et de respect à votre égard.

A mes tous amis

A mes amies Mariame Sifallah et Najwa Chaydy avec qui j'ai passé de très agréables moments, je dédie ce travail en guise de gratitude pour leur grand soutien et pour leur affection sans limite.

A Ghizlane

A mon binôme Ghizlane pour son fort esprit d'équipe et son sens de l'humour qui rend le travail agréable et moins stressant.

Amal

Remerciements

En premier lieu, nous tenons à présenter nos sincères remerciements à M. Mohammed Boutellaka, chef de service actuariat et surveillance de portefeuille ainsi que maître de stage, pour nous avoir fait confiance dans la réalisation de ce projet. Le travail sous sa direction nous a permis d'élargir nos horizons et d'affiner nos pistes de réflexion. Par la même occasion, nous gratifions M. Ikkal Chahdi Ouazzani et Mme Moufaddala Ballouta, membres de l'équipe actuariat, pour leur attitude agréablement amicale.

Nous adressons également nos profonds remerciements à M. Hassan Alami, pour avoir bien voulu partager son immense expérience avec nous débutantes, d'avoir chaperonné notre travail depuis son début et d'avoir toujours répondu présent quand nous avons besoin d'un avis d'expert.

Nos grands remerciements vont vers M. Fouad Marri, notre professeur et encadrant interne, pour son grand savoir faire qu'il nous a volontiers communiqué. Nous le remercions également pour sa patience, sa modestie et sa disponibilité.

Finalement, nous tenons à rendre hommage à l'ensemble du corps professoral de l'INSEA qui, durant nos trois ans de parcours à l'institut, a œuvré à maintenir une excellente qualité de formation et à nous inculquer les grandes valeurs du professionnalisme.

Résumé

Le présent rapport constitue la trace écrite du travail que nous avons réalisé dans le cadre de notre projet de fin d'études. Ce dernier fût effectué au sein du groupe mutualiste MAMDA-MCMA et porte sur l'estimation de la provision pour sinistres à payer pour le portefeuille AT.

Le commencement était une étude documentaire effectuée principalement au centre de documentation de l'entreprise d'accueil. Par la suite, nous avons procédé à l'analyse des méthodes de provisionnement les plus populaires afin de se familiariser avec quelques principes et concepts de base. Cette phase préliminaire fut indispensable au bon avancement du projet dans la mesure où elle nous a donné une idée assez précise sur les techniques de provisionnement recommandées.

La sélection de l'outil informatique adéquat à l'étude fut notre seconde préoccupation. Le choix tomba sur le progiciel SAS pour sa grande capacité d'adaptation à toute sorte de problèmes de calcul matriciel ou d'étude statistique par l'implémentation de diverses procédures. Son langage de programmation aussi simple que complet laisse une grande marge de manœuvre et s'avère particulièrement adapté aux problèmes de provisionnement. Cela dit, quelques failles relatives à la version installée du logiciel seront comblées par l'usage d'autres logiciels notamment le R ou Excel...

La réalisation de ce projet s'est fondamentalement basée sur la manipulation des triangles de liquidation. La méthode de Chain Ladder Standard sera la technique par défaut et se retrouvera assez souvent dans diverses parties de ce travail. Par ailleurs, l'accent sera mis sur les méthodes stochastiques et sur le calcul de risque relatif à chaque estimation de provision. En particulier, quelques techniques statistiques de simulation seront abordées pour plus de vigilance dans les calculs.

Mots clés : Provisionnement stochastique, Bootstrap, Modèle linéaire généralisé, Chain Ladder, Best estimate.

Abstract

This report is the written record of the work we have done in the framework of our project graduation. This was done within the group MAMDA-MCMA and it aims to estimate the provision for outstanding claims of the branch work accident.

The beginning was a literature review conducted mainly in the documentation center of the host company. Subsequently, we analyzed the most popular provisioning methods in order to become familiar with the basic concepts. This preliminary phase was essential to the smooth progress of the project so far as it gave us a fair idea about the recommended provisioning techniques.

Selecting the suitable software tool for the study was our second concern. The choice fell upon the SAS software package for its ability to adapt to any kind of matrix algebra problems or statistical study by the implementation of various procedures. Its programming language is simple and complete so it allows flexibility and is particularly suited to the problems of provisioning. However, some faults on the installed software will be filled by the use of other software such as Excel or R...

The achievement of this project is fundamentally based on the manipulation of liquidation triangle. Chain Ladder Standard Method will be the technical default and will be found quite often in various parts of this work. Moreover, the focus will be on stochastic methods and calculation of relative risk estimate for each provision. Specifically, some statistical techniques for simulation will be applied in order to insure more vigilance in the calculations.

Key words: stochastic provisioning, Bootstrapping, General linear model, Chain Ladder, Best estimate.

[Tapez le titre du document]

Table des matières

[Tapez le titre du document]

Liste des figures[Tapez le titre du document]

Liste des figures

Liste des tableaux

Liste des tableaux

Liste des abréviations

Abréviation	Désignation
IBNR	Incured But Not Reported
RBNP	Reported But Not Paid
PSAP	Provisions pour Sinistres A Payer
AT	Accident de Travail
MAMDA	Mutuelle Agricole Marocaine D'Assurances
MCMA	Mutuelle Centrale Marocaine d'Assurances
UMAM	Union Marocaine des Assurances Mutuelles
SP	Sinistres Payés
SPC	Sinistres Payés Cumulés
SAP	Sinistres A Payer
PA	Primes Acquises
MCO	Moindres Carrés Ordinaires
MACIF	Mutuelle Assurance des Commerçants et Industriels de France
OIT	Organisation Internationale du Travail
UMAM	Union Marocaine des Assurances Mutuelles
MAEM	Mutuelle Assurances des Enseignants du Maroc
GASup	Groupement d'Assurance du Supérieur
SAM	Sociétés des Assurances Mutuelles
DAPS	Direction des Assurances et de Prévoyance Sociale
D/D	Dossier par dossier
CM	Coût moyen
GLM	Modèle linéaire généralisé

Introduction générale

La génération de revenu et la production et pour les entreprises pratiquant une activité d'assurance sont très particulières. Quand une société commerciale ordinaire se base sur le coût de revient exact pour tarifier son produit en le majorant simplement d'une certaine marge de bénéfice, l'assureur, quant à lui, se heurte à l'ignorance totale ou partielle de la charge qu'entraînera la vente d'un contrat à l'un de ses assurés. Cette incertitude face au futur, le risque de faire une tarification déficitaire ou encore les conséquences périlleuses d'une sous-estimation des provisions font de l'assurance l'un des domaines les plus sévèrement réglementés. Ainsi, une compagnie d'assurances est tenue de contrôler en permanence ses bases tarifaires et d'augmenter ses provisions techniques dès que celles-ci s'avèrent insuffisantes.

Les provisions techniques recouvrent l'ensemble des charges et paiements à prévoir en cas d'une éventuelle sinistralité. D'après l'exigence du Ministère des finances, tous les organismes assureurs doivent constituer des réserves selon un schéma réglementaire unifié. En vue de réunir les données requises à l'ensemble de ces estimations, la tenue d'une écriture comptable correcte fait également partie des mesures organisationnelles exigées d'une entreprise d'assurances. Etant des dettes, les provisions techniques sont inscrites au passif du bilan. Plus encore, elles constituent plus que sa moitié. Ainsi, une petite erreur d'estimation du montant des provisions entraînerait une considérable variation du résultat de l'entreprise. Par ailleurs, la solvabilité de cette dernière serait autrement appréciée. Conséquemment, une grande vigilance est fortement préconisée pour ce type de calcul. Les provisions techniques font naturellement l'objet de plusieurs placements. Par conséquent, l'entreprise doit assurer à chaque moment une certaine liquidité pour un maximum de sécurité.

Le présent projet se donne pour objectif de cerner avec précision la valeur de la PSAP au titre de l'exercice 2007 pour la branche AT du groupe MAMDA-MCMA. Cela dit, le fait de se référer à l'année 2007 montre bien que le résultat obtenu n'est pas un but en soit. En effet, il ne permet que d'avoir une plateforme chiffrée facilitant l'exploitation des diverses méthodes et techniques de provisionnement dédiées à l'assurance non-vie. Travailler sur la réserve d'une année révolue permet également de vérifier la pertinence des estimations obtenues suite à leur comparaison aux résultats effectifs. Pour récapituler, nous tentons à travers ce projet de présenter diverses méthodes de provisionnement. Nous focalisons particulièrement nos efforts sur la simplification des techniques souvent lourdes en formules mathématiques complexes. L'accent est également mis sur la compréhension de la philosophie se tenant derrière la construction de chaque modèle. En résumé, notre travail vise à laisser une trace écrite simplifiée pour l'application de chacune des techniques sélectionnées et à les accompagner du code SAS donnant selon chaque méthode le calcul de la réserve finale.

La constitution et l'évaluation des provisions techniques fait l'objet d'une grande intervention de la part de l'autorité de contrôle des assurances. La réglementation est précise sur ce point et fournit selon un schéma concis un ensemble de méthodes enchaînées aboutissant à la valeur de la réserve à retenir. Cependant, ce schéma réglementaire de provisionnement s'inscrit dans un cadre préventif et par conséquent donne lieu à une surestimation de la provision à constituer.

L'approche déterministe pour le calcul des provisions techniques revêt un aspect comptable. Principalement fondée sur les éléments du triangle de liquidation, cette approche consiste en l'application de quelques opérations arithmétiques directes sur les données utilisées. Le résultat est une valeur précise de la réserve qui varie selon la méthode suivie. Il se trouve que la mieux connue et la plus utilisée des techniques déterministes est Chain Ladder Standard. Cette dernière se base sur quelques hypothèses faites à propos des éléments sur lesquels elle s'opère. L'algorithme qui va avec cette technique est certes simple, mais il ne permet pas d'avoir une idée sur l'ampleur de l'incertitude inhérente à telle ou telle estimation.

Introduction générale

Ce défaut majeur n'est pas exclusif à Chain Ladder : il est présent dès qu'on parle de méthode non probabiliste. De ce fait, le recours aux méthodes stochastiques est impératif.

L'usage des calculs de probabilités pour décrire le comportement des variables réserve et incréments de paiements relève des compétences de l'actuaire et fait l'objet des techniques stochastiques dédiées au provisionnement. Que l'on soit en présence de méthodes paramétriques ou non-paramétriques, l'objectif est de cadrer la valeur prédite d'une certaine variable, notamment la réserve, entre les bornes d'un intervalle de confiance. Cela se fait après calcul de divers types de risques d'estimation. Notons par ailleurs que l'approche stochastique est reconnue pour être plutôt complète en ce sens qu'elle permet d'obtenir des informations sur la distribution d'un état futur.

Chaque technique stochastique a ses propres spécificités. Le modèle de Mack tient son importance du fait qu'il n'est autre que la version probabilisée de la méthode Chain Ladder. La technique Munich Chain Ladder se caractérise quant à elle par l'inclusion de l'élément exposition au risque lors du calcul de la provision. La modélisation GLM se distingue par sa flexibilité et par son aptitude à décrire le comportement aléatoire des paramètres d'intérêt du modèle. Pour ce qui est de la technique du Bootstrap dite méthode de rééchantillonnage, elle se couple aux diverses méthodes précitées pour créer un supplément d'information à partir des données déjà disponibles.

Ainsi, l'objectif de notre travail est de trouver, avec le plus possible de précision, la valeur estimative de la PSAP. Dans cette optique, une première partie est dédiée à la présentation des différents éléments de la problématique. La seconde est réservée aux techniques déterministes, allant des méthodes réglementaires recommandées par l'autorité de tutelle jusqu'aux méthodes de ratios ou encore les variantes de Chain Ladder. La troisième partie retrace quant à elle les diverses étapes de construction et de mise en œuvre des modèles stochastiques utilisés :

- Le modèle de Mack.
- Le modèle de Munich Chain Ladder.
- La technique du Bootstrap.
- La modélisation GLM.

Bien entendu, chacune des deux dernières parties se termine sur une comparaison des résultats issus des différentes méthodes afin de trancher sur la meilleure estimation possible de la provision. Finalement, et en guise de récapitulatif, une conclusion générale vient résumer les principaux résultats et constats établis tout au long du travail.

Partie I. Etude préliminaire

Cette première partie, qui est l'étude préliminaire du projet est scindée en trois chapitres.

Le premier chapitre présente le mutualisme comme philosophie d'assurance avant de donner un bref aperçu sur l'historique du secteur mutualiste au Maroc. Ce chapitre fournit également une présentation concise de l'organisme d'accueil.

Le second chapitre explicite les notions de base relatives au provisionnement tout en citant, selon chaque type d'assurance, quelques provisions devant être constituées par l'assureur. Il donne également une justification chiffrée du choix de l'étude des PSAP.

Le dernier chapitre de cette partie, quant à lui, définit la branche d'assurance AT sous un angle législatif. Par la suite, il définit et décrit les différents éléments de la base de données étudiée dans le cadre de ce projet.

Introduction

« Nous avons une convergence de valeurs », avança Gérard Andreck, président du conseil d'administration de la Macif, dans le cadre de la conclusion d'un accord de partenariat entre la Macif et le groupe MAMDA-MCMA. Il rajouta : « Nous ne nous marions pas avec n'importe qui ! Nous choisissons les partenaires qui partagent notre projet mutualiste. C'est primordial. Et la MAMDA-MCMA est un partenaire idéal ». Venant du président d'une fondation dont le chiffre d'affaires dépasse de plus de 10 fois celui du marché marocain dans son ensemble, un tel propos pèse lourd et témoigne en la faveur du groupe mutualiste marocain.

La MAMDA-MCMA se trouve au cœur de plusieurs grands projets de développement nationaux tel que le plan Maroc vert. Le groupe mutualiste ne cesse d'élargir sa gamme de produits et d'apporter des solutions innovantes qui s'adaptent aux besoins grandissant de ses sociétaires, comme ce fût le cas du tout récent « multirisque climatique », lancé depuis la fin de l'année 2011.

Cependant, le groupe MAMDA-MCMA ne connut pas que des gloires. En effet, la fin des années 60 fût particulièrement pénible pour la compagnie qui était au bord de la faillite avec un chiffre d'affaires de 7 millions de Dhs et des réserves qui ne dépassent guère les 10 millions de Dhs. Suite à cette faille dans le système, les professionnels du groupe ont veillé à booster ses performances. C'est dans ce sens que la compagnie fait appel aux connaissances et compétences actuarielles.

Comme toute autre compagnie d'assurance, la MAMDA-MCMA se soucie de l'exactitude de ses chiffres et prévisions. En particulier, l'art de la constitution des provisions est un pilier majeur sur lequel reposent une série d'implications relatives à la solvabilité de la société. De ce fait, le calcul des réserves capte une attention particulière et nécessite de réunir des compétences mathématiques avancées, d'avoir une parfaite connaissance de la sinistralité passée grâce aux pièces justificatives et documents comptables et de disposer d'un ensemble d'outils informatiques et progiciels facilitant l'application des plus complexes techniques et méthodes.

Dans la diversité des produits assuranciers proposés par le groupe MAMDA-MCMA, l'assurance AT gagne de plus en plus d'ampleur. Effectivement, selon un rapport publié par l'OIT en fin 2011, environ 317 millions d'employés par an sont blessés sur leur lieu de travail, soit 850.000 accidents quotidiens nécessitant inévitablement arrêts de travail et remboursements. Ainsi, l'effectif croissant des victimes de ce genre d'accidents et le caractère obligatoire de l'assurance AT font d'elle une branche de grande envergure.

Chapitre I. Introduction au secteur et à l'organisme d'accueil

I. Aperçu sur le secteur mutualiste:

La société mutuelle est une personne morale de droit privé à but non lucratif. Elle s'assimile aux sociétés coopératives et aux associations dans la mesure où ses clients voire même ses employés ne sont autres que ses sociétaires. Ainsi, une société mutuelle n'est ni une entreprise commerciale ni un organisme caritatif.

Ce type de sociétés fonctionne grâce à la mise en commun des moyens de chacun des associés en vue de faire face aux aléas qui les guettent. Le système de collecte ne prend pas en considération le risque individuel. Cela dit, l'équilibre financier est important. Ainsi la politique tarifaire d'une société mutuelle diffère de celle d'une société commerciale qui se base sur le rapport « sinistre/prime ».

Le système mutualiste a fait son apparition au Maroc pour la première fois avec la venue du protectorat. La raison du recours à ce système fût le besoin d'assurance ressenti par le personnel de l'administration du Protectorat français. La santé a été le premier domaine couvert par le mutualisme. Par la suite, d'autres domaines à risque furent pris en charge, notamment aux niveaux agricole et industriel.

II. Présentation du groupe MAMDA-MCMA

II.1 Création et description de l'entreprise

Organisées sous forme de caisses régionales, les mutuelles agricoles ont fait leur apparition au Maroc en 1920. Nées sous le protectorat, elles avaient alors pour mission de répondre aux besoins des seuls colons. Elles furent reprises et fédérées par les agriculteurs marocains en 1963 sous la bannière de la toute nouvelle MAMDA.

En 1968, naît la MCMA dont l'objet est la prise en charge des risques non agricoles des agriculteurs et la promotion du secteur mutualiste dans toutes les couches de la société marocaine. La MCMA fédère des mutuelles et des sociétaires appartenant à toutes les professions. Les enseignants du secondaire (MAEM) et du supérieur (GASUP) forment la colonne vertébrale de la MCMA.

MAMDA et MCMA sont deux mutuelles juridiquement et financièrement indépendantes mais qui partagent management, personnel, locaux et ressources commerciales. MAMDA et MCMA devraient s'unir sous le nom d'UMAM tout en conservant leur autonomie juridique et commerciale. La croissance de la nouvelle entité devrait reposer sur la l'adhésion de nouvelles mutuelles dont elle encouragerait la création.

Le Groupe MAMDA-MCMA noue des partenariats internationaux à la fois pour améliorer la qualité de ses produits et pour étendre son bassin de clientèle aux Marocains résidant à l'étranger et aux expatriés étrangers résidant au Maroc.

II.2 Mission

La MAMDA et la MCMA sont les deux principales et plus anciennes mutuelles d'assurance marocaines. La MAMDA a une vocation rurale tandis que la MCMA est plus généraliste. Juridiquement indépendantes, elles partagent les mêmes valeurs de l'économie sociale: solidarité, justice et performance financière. Historiquement liées par une étroite collaboration technique, humaine et commerciale, elles ont reçu l'autorisation de la Direction des assurances de se constituer en Union de sociétés d'assurances mutuelles (SAM) susceptible d'abriter d'autres corps professionnels. Fédérées par la future Union, les SAM conserveront, entre autre, leur autonomie commerciale.

II.3 Produits

Les produits du groupe sont à la fois des produits d'assurances Agricoles et des produits d'assurances non Agricoles. La Mutuelle Agricole couvre les risques suivants :

- Accident de Travail et Maladies Professionnelles.
- Grêle.
- Incendie Exploitation Agricole.
- Incendie Matériel Agricole.
- Incendie Pailles et Fourrages.
- Incendie Récolte sur pied.
- Incendie Explosions.
- Individuel Accident Agricole.
- Mortalité de bétail.
- Multirisque Automobile.
- RC Agricole.
- Vol.
- Multirisque climatique.

La MAMDA gère pour le compte de l'Etat le programme de garantie des céréales contre les effets de la sécheresse.

Afin de répondre aux besoins du monde non agricole en matière d'assurances, la MCMA offre à ses sociétaires une gamme complète de produits concernant l'ensemble des risques des particuliers et des entreprises à savoir :

PRODUITS DOMMAGES

- Accident de Travail et Maladies Professionnelles.
- Individuel Accident.
- Multirisque des commerçants Artisans.
- Multirisque des professions Libérales.
- Multirisque Habitation.
- RC Chasse.
- RC Exploitation.
- RC Scolaire.
- Tous Risques Chantier.
- Transport de Marchandises.
- Vol.
- Multirisque Automobile.
- Multirisque Hôtel.
- Incendie Explosion.

PRODUITS VIE

- Maladie-Maternité.
- Assurance Décès.
- Compte Avenir Jeune.
- Epargne Retraite Mutuelle.

Chapitre II. Les provisions techniques

I. Définition générale des provisions techniques

Les provisions techniques est la désignation du montant que l'assureur doit mettre de côté en vue de faire face aux charges éventuelles de la sinistralité. Elles servent également à anticiper les prestations que l'assureur doit verser aux bénéficiaires suivant les termes du contrat. De façon générale, il s'agit de la somme affectée par l'entreprise d'assurances à la couverture d'une charge ou d'une perte virtuelle, future ou éventuelle relative à son engagement envers les assurés.

Le terme technique permet de faire la distinction des autres provisions comme les provisions pour risques et charges ou encore les provisions pour dépréciation.

Les règles de calcul de ces provisions sont diverses:

- Les règles légales ou réglementaires : tendent à définir le montant à provisionner selon les attentes du législateur. Elles sont puisées dans les textes et publications juridiques du Ministère de l'Economie et des Finances.
- Des règles professionnelles, relevant du « bon usage ».
- Des méthodes mathématiques et modélisations probabilistes relevant des compétences actuarielles.

Très vaguement, pour le cas de l'assurance de personnes à déroulement court ou bien pour le cas de l'assurance non-vie, les provisions sont tributaires du décalage possible entre la survenance du fait dommageable, générateur du paiement de l'indemnité et le règlement effectif de cette indemnité. Nous avons donc deux types principaux de provisions:

- Provisions pour sinistres connus ; reportés mais pas encore payés (RBNP).
- Provisions pour sinistres inconnus; survenus mais pas encore déclarés qu'on nomme tardifs (IBNR).

II. Exemples de provisions pour quelques types d'opérations d'assurances

II.1 Quelques opérations d'assurance-vie

Ici sont regroupées les provisions propres aux catégories d'opérations d'assurances suivantes :

- a. Vie et décès.
- b. Nuptialité et natalité.
- c. Capitalisation.
- d. Opérations ayant pour objet l'acquisition d'immeubles au moyen de la constitution de rentes viagères.
- e. Assurances liées à des fonds d'investissement.
- f. Opérations faisant appel à l'épargne dans le but de réunir les sommes versées par les assurés en vue de la capitalisation en commun.

Pour les opérations d'assurances précitées, le législateur a prévu la constitution des provisions suivantes :

- Provisions mathématiques : Il s'agit de la différence à la date d'inventaire entre les valeurs actuelles des engagements respectivement pris par l'assureur et par les assurés.
- Provision de gestion : Provision destinée à couvrir les charges de gestion futures des contrats non couvertes par ailleurs.
- Provision pour capitaux et rentes à payer : C'est la valeur des capitaux et rentes échus et restant à payer à la date de l'inventaire.

- Provision pour participation aux bénéficiaires : Il s'agit d'une provision constituée pour enregistrer l'engagement de l'entreprise d'assurance envers les bénéficiaires des contrats lorsque les montants dus au titre des bénéficiaires n'ont pas encore été versés au compte de l'assuré.
- Provision pour fluctuations de sinistralité : Cette provision permet de faire face aux charges que l'entreprise pourrait être amenée à supporter en cas de survenance d'un ou plusieurs sinistres exceptionnels.
- Provision de capitalisation : Le montant destiné à pallier à la dépréciation ou à la diminution du revenu des éléments inscrits à l'actif.
- Provision pour aléas financiers : Le montant mis de côté pour faire face à la diminution du rendement de l'actif.
- Provision pour risque d'exigibilité : Ces provisions servent, en cas de dépréciation des valeurs de l'actif, à faire face aux engagements techniques.

II.2 Opération d'assurances AT

Les provisions exigées pour les opérations d'assurances de type accident de travail sont comme suit :

- Provisions mathématiques.
- Provision pour arrérages échus : Il s'agit, à la date de l'inventaire, du montant de la partie de la rente due, échue et restant à payer.
- Provision pour primes non acquises : Ici, il s'agit de rendre compte, pour les contrats à primes émises d'avance, de la partie des primes correspondant à la période comprise entre la date de l'inventaire et celle de la prochaine échéance ou, à défaut, du terme du contrat.
- Provision pour risques en cours : Cette provision sert, pour la période allant de la date d'inventaire à celle de la prochaine échéance ou, à défaut, du terme du contrat, à couvrir les sinistres et les charges complétant ainsi le montant non pris en compte par la provision pour primes non acquises.
- Provisions pour sinistres à payer : Il s'agit des provisions servant à faire face aux sinistres dont le règlement des charges y afférentes n'a pas encore été effectué. Plus précisément, c'est le montant des dépenses pour sinistres non réglés ou celui des dépenses pour sinistres réglés restant à payer à la date de l'inventaire.
- Provisions pour appareils de prothèse : C'est le montant destiné à indemniser l'assuré en cas d'achat ou même de renouvellement d'appareils de prothèse.
- Provision pour fluctuations de sinistralité.
- Provision de capitalisation.
- Provision pour aléas financiers.
- Provision pour risque d'exigibilité.

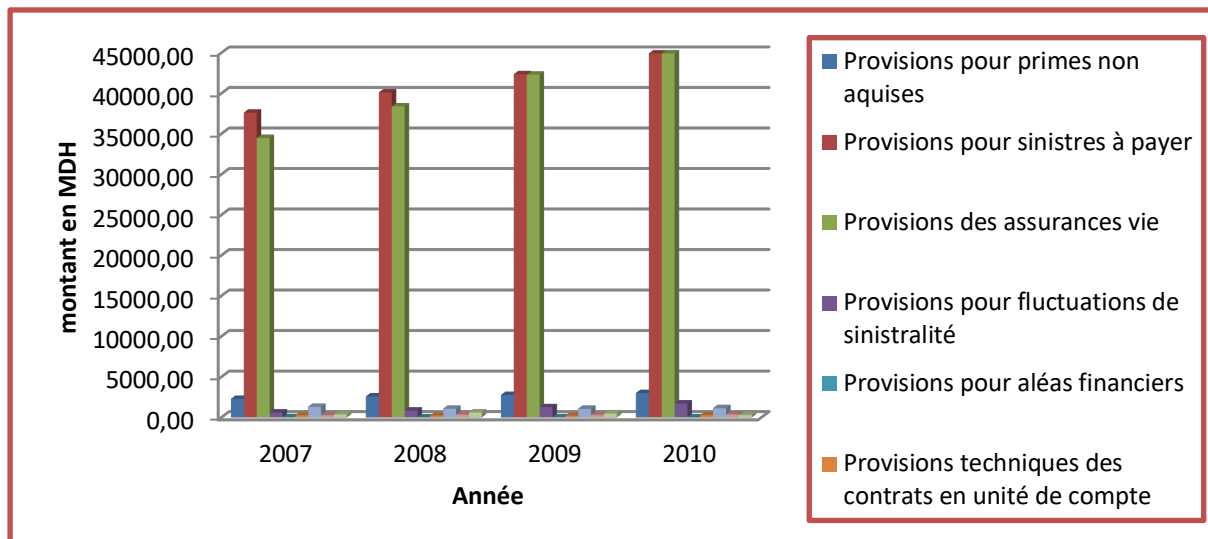
II.3 Justification statistique du choix des PSAP

Etant donné la diversité des provisions techniques, choisir de calculer l'une d'entre elles doit être dûment justifié. Pour cela, nous avons pensé à comparer le volume de chacune de ces provisions pour la période s'étalant de l'année 2007 à l'année 2010.

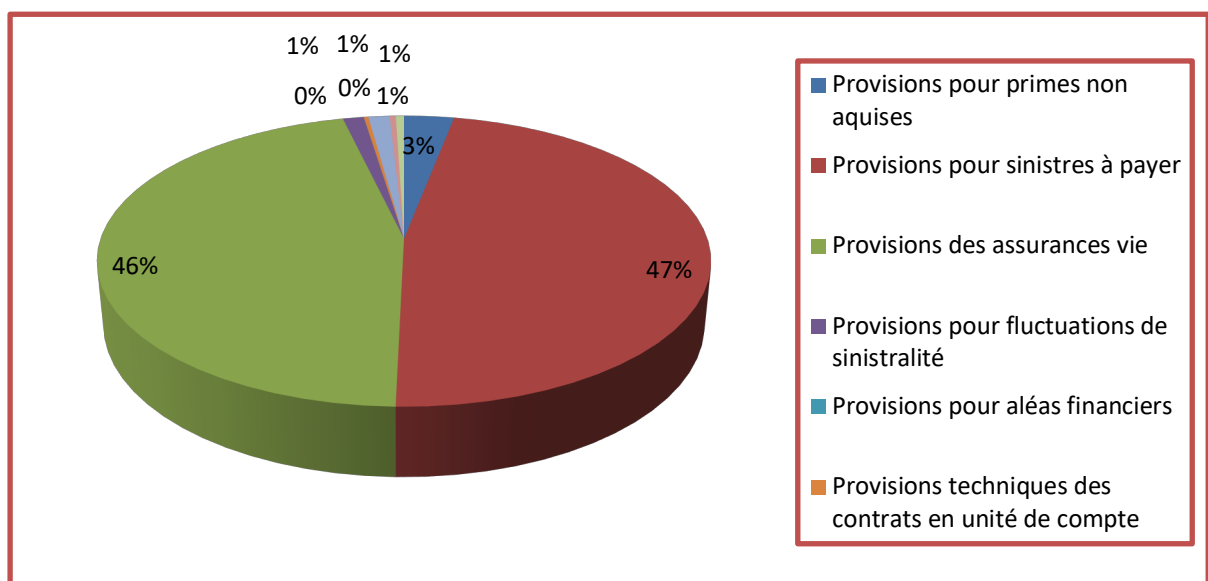
Provisions techniques par type (en MDH)	2007	2008	2009	2010
Provisions pour primes non acquises	2283,96	2593,79	2767,50	3004,73
Provisions pour sinistres à payer	37612,80	40105,45	42350,85	44912,44
Provisions des assurances vie	34475,63	38376,50	42298,63	44911,28
Provisions pour fluctuations de sinistralité	606,74	844,64	1262,14	1721,49
Provisions pour aléas financiers	9,71	4,45	14,01	0,00
Provisions techniques des contrats en unité de compte	251,71	226,55	207,56	267,67

Provision pour participation aux bénéfices	1291,02	1055,65	1052,54	1129,12
Autres provisions techniques	266,64	357,06	312,45	374,79
Provisions techniques sur placements	364,51	598,39	476,30	296,79
Total	77162,72	84162,48	90741,98	96618,31

Pour mieux sonder notre comparaison, nous passons par un graphe représentatif des montants des différentes provisions et ce, pour chacune des quatre années dont les données sont disponibles.



Un bref coup d'œil sur le diagramme ci-dessus laisse entrevoir la prédominance triviale des PSAP sur le reste des provisions. En effet, le montant des PSAP à lui seul égalise celui de l'ensemble des provisions pour l'assurance-vie. Le reste des provisions, à savoir les provisions pour primes non acquises, les provisions pour fluctuations de sinistralité, les provisions pour aléas financiers et les provisions techniques des contrats en unité de compte sont clairement négligeables en termes d'importance devant le volume des PSAP.



Le diagramme en secteur de la moyenne sur les années 2007-2010 des montants des provisions nous confirme l'idée déjà établie sur l'importance cruciale des PSAP. Ainsi, ces dernières détiennent 47% du montant global des provisions de tout le marché. Le reste des

provisions techniques propres aux opérations de type assurances non-vie sont infinitésimales et, groupées ensemble, n'arrivent qu'à 7% du total.

Ainsi, vu leur indéniable importance, nous consacrons notre projet à l'étude minutieuse des PSAP.

Chapitre III. Eléments de l'étude et problématique

I. Introduction à la branche AT

Les accidents survenus par le fait ou à l'occasion du travail donnent droit au profit de la victime (salariée ou travaillant à quelque titre ou en quelque lieu que ce soit, même si l'employeur n'exerçait pas sa profession dans un but lucratif) ou de ses représentants à une indemnité à la charge de l'entreprise ou de l'employeur. Pour l'identification d'un accident de travail, il suffit de constater que le travail a été acteur dans le dommage corporel même si l'accident en question n'a fait qu'aggraver ou mettre en mouvement les effets des forces de la nature. Par dérogation, l'employeur ou l'assureur peut rapporter la preuve d'une disposition pathologique de la victime. L'AT comprend également les accidents de trajet entre le domicile et le lieu du travail en plus des accidents survenus au cours d'un déplacement professionnel ou se qu'on appelle accidents de mission.

Récemment, la législation marocaine a clairement évolué garantissant ainsi à l'employé tous ses droits et en lui assurant un maximum de sécurité lors de l'exercice de son activité professionnelle. Les textes réglementant le travail sont stricts : ils détaillent désormais tous les points impliqués à savoir l'information de l'employeur, la déclaration par l'employeur, son préposé ou la victime de l'accident auprès des autorités compétentes, l'enquête judiciaire menée par le juge de première instance à l'issue de laquelle une décision statuant sur les indemnités dues par l'employeur ou son assureur au profit de la victime.

Etant donc une branche d'assurance de grande envergure, nous ferons de l'AT l'objet de notre projet. Plus précisément, on essaiera de calculer les provisions pour sinistres à payer pour le produit accident de travail à partir des données du groupe MAMDA-MCMA.

II. Présentation des données utilisées

II.1 Descriptif de la base de données

Les données dont on dispose et qui feront la matière première de toutes les techniques de calcul de provisions sont principalement les éléments du triangle de liquidation. Ce dernier comprend, dans sa version initiale, l'ensemble des charges et des paiements effectués relativement aux sinistres déclarés dans la période s'étalant de 1997 jusqu'à 2007. Le déroulement total de la branche est supposé de 11 ans. Cette durée n'est pas fixée à l'avance mais on l'a déduite à partir des données. En effet, on constate que 11 années est une durée suffisante pour la liquidation de la grande majorité des sinistres.

Pour une meilleure compréhension des données, nous donnons la signification de chacun des montants, éléments du tableau :

- ❑ SP: il s'agit des paiements des sinistres totalement ou partiellement réglés. Ainsi, SP comprend également le montant versé dans le cas des sinistres non terminés. Les SPC ou sinistres payés cumulés sont, comme leur nom l'indique, pour une même année de survenance, le cumul des SP année par année.
- ❑ SAP: c'est le montant estimatif des charges pour sinistres déclarés non encore payés ou partiellement payés.

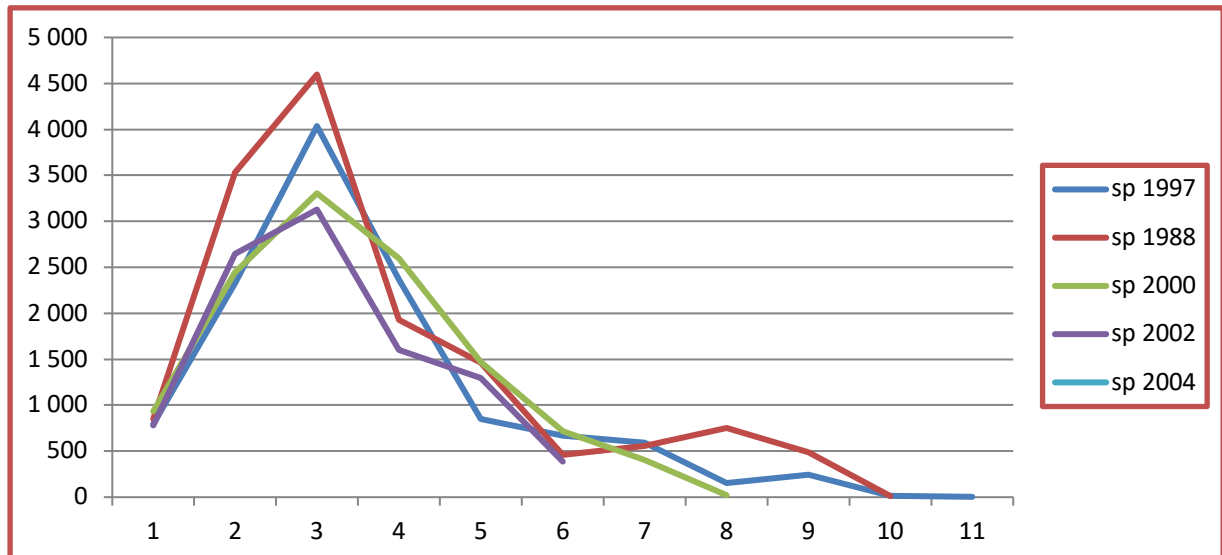
Le tableau donne également le montant des PA pour chaque année de déclaration. De ce fait, le tableau fourni est une imbrication des deux triangles de liquidation classiques:

- ❑ Le triangle des paiements.
- ❑ Le triangle des charges.

		1 997	1 998	1 999	2 000	2 001	2 002	2 003	2 004	2 005	2 006	2 007	P.A.
1 997	SPC	xxxxx	792	3 119	7 155	9 522	10 370	11 036	11 631	11 783	12 029	12 041	
	SP	792	2 327	4 036	2 367	848	666	595	152	246	12	4	
	S.A.P	6 957	10 011	7 887	3 975	3 004	2 240	1 500	1 298	963	798	279	
	TOTAL	7 749	13 130	15 042	13 497	13 374	13 276	13 131	13 081	12 992	12 839	12 324	12 860
1 998	SPC		xxxxx	847	4 380	8 977	10 908	12 370	12 831	13 392	14 143	14 631	
	SP		847	3 533	4 597	1 931	1 462	461	561	751	488	12	
	S.A.P		9 211	12 081	6 715	4 816	3 300	2 242	1 847	1 793	1 444	1 252	
	TOTAL		10 058	16 461	15 692	15 724	15 670	15 073	15 239	15 936	16 075	15 895	10 026
1 999	SPC			xxxxx	708	3 253	6 300	8 179	9 718	10 287	10 518	10 662	
	SP			708	2 545	3 047	1 879	1 539	569	231	144	7	
	S.A.P			8 944	9 922	6 636	5 050	3 598	2 803	1 786	1 053	952	
	TOTAL			9 652	13 175	12 936	13 229	13 316	13 090	12 304	11 715	11 621	14 138
2 000	SPC				xxxxx	933	3 378	6 683	9 282	10 753	11 473	11 880	
	SP				933	2 445	3 305	2 599	1 471	720	407	22	
	S.A.P				7 832	8 680	5 986	4 319	2 891	2 101	1 147	963	
	TOTAL				8 765	12 058	12 669	13 601	13 644	13 574	13 027	12 865	12 086
2 001	SPC					xxxxx	783	3 291	5 852	7 393	8 202	8 653	
	SP					783	2 508	2 561	1 541	809	451	479	
	S.A.P					5 840	6 622	4 025	3 431	2 402	1 761	1 124	
	TOTAL					6 623	9 913	9 877	10 824	10 604	10 414	10 256	12 856
2 002	SPC						xxxxx	783	3 429	6 557	8 159	9 458	
	SP						783	2 646	3 128	1 602	1 299	387	
	S.A.P						5 812	7 769	5 239	4 012	3 718	2 955	
	TOTAL						6 595	11 198	11 796	12 171	13 176	12 800	11 580
2 003	SPC							xxxxx	816	6 227	12 074	15 372	
	SP							816	5 411	5 847	3 298	1 242	
	S.A.P							12 186	16 133	10 321	10 173	9 786	
	TOTAL							13 002	22 360	22 395	25 545	26 400	18 483
2 004	SPC								xxxxx	1 156	5 154	11 534	
	SP								1 156	3 998	6 380	4 396	
	S.A.P								10 104	14 541	10 759	6 878	
	TOTAL								11 260	19 695	22 293	22 808	19 531
2 005	SPC									xxxxx	903	5 007	
	SP									903	4 104	4 229	
	S.A.P									9 850	15 410	9 573	
	TOTAL									10 753	20 417	18 809	15 652
2 006	SPC										xxxxx	977	
	SP										977	2 781	
	S.A.P										11 740	12 579	
	TOTAL										12 717	16 337	18 815
2 007	SPC											xxxxx	
	SP											1 122	
	S.A.P											22 254	
	TOTAL											23 376	30 442

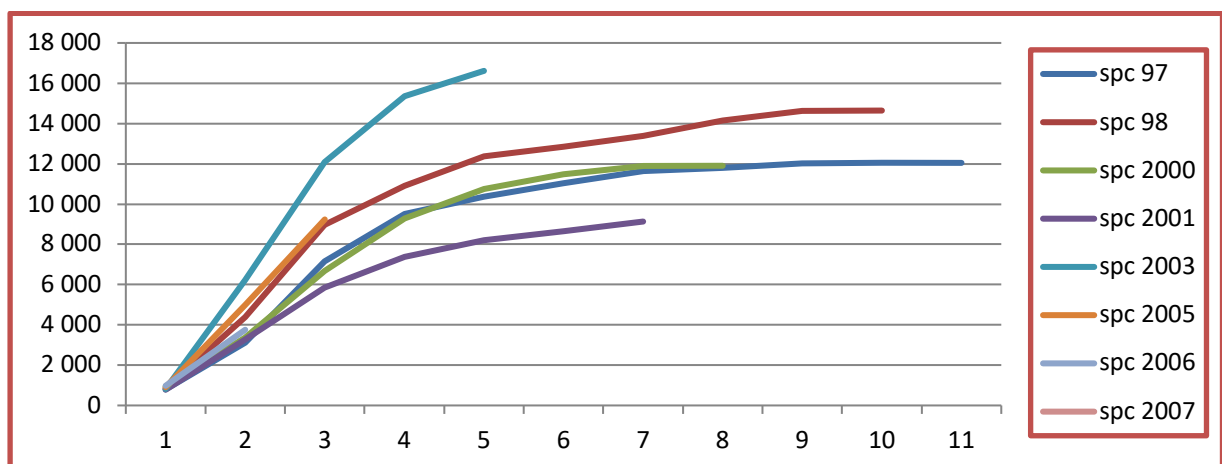
2 006	977	2 781
2 007	1 122	

Le diagramme ci-dessous est la représentation graphique simultanée d'une multitude de courbes. Chacune d'elles se réfère à une année de survenance. Ainsi, nous retrouvons sur l'ordonnée le montant des paiements de sinistres et sur l'abscisse le délai de règlement correspondant.



De manière assez directe, le constat qui s'impose est que toutes les courbes ont une allure très semblable. En effet, pour les deux premières années de déroulement, le montant de règlement est ascendant. Le paiement des sinistres connaît généralement son apogée durant la 3^{ème} année d'après la survenance. Juste après, le montant chute brusquement avant de s'atténuer presque pour les années qui suivent la 5^{ème}.

Nous faisons les mêmes observations en raisonnant en termes de paiements cumulés :



Il est ainsi clair que le cumul des paiements s'accroît de façon spectaculaire durant les trois premières années. Les paiements qui se font par la suite sont moins volumineux ; cela se voit à la petite pente que prend la courbe une fois la 3^{ème} année dépassée.

Cette même analyse peut se faire à l'aide de la cadence de règlement ou des facteurs de développement mais avant d'entamer l'analyse à l'aide de ces facteurs, on se doit de définir les deux concepts qui s'y rapportent.

En effet, les facteurs de développement sont des ratios retraçant l'ampleur de l'évolution des paiements effectués entre deux délais consécutifs, toutes années de survenance confondues. Concrètement, le facteur de développement entre les deux délais j et $j+1$ est:

$$F_j = (\sum_{i=0}^{n-j-1} C_{i,j+1}) / (\sum_{i=0}^{n-j-1} C_{ij}), \quad j = 0, \dots, n-1 \text{ avec } n=10 \text{ pour notre cas.}$$

C_{ij} Étant le paiement des sinistres survenus l'année i et qui est effectué au délai j , c'est-à-dire en l'année $i+j$.

Les cadences cumulées de règlement sont, quant à elles, une autre forme de ratios beaucoup plus facile à interpréter. Il s'agit exactement d'un rapport entre le cumul des paiements jusqu'à un délai j effectués pour les sinistres survenus au cours d'une même année i sur la charge ultime correspondant à cette même année d'origine. La charge ultime pour une année d'origine i correspond à la totalité des paiements de sinistres survenus en i , y compris les paiements dont on ne dispose pas et que nous devons estimer par la suite.

Si l'on note C_{in} la charge ultime correspondant à l'année de survenance i , alors on écrit $C_{in} = C_{ij} * F_j * F_{j+1} * \dots * F_{n-1}$ pour $j=n-i \dots n-1$. Ce point sera détaillé par la suite quand on abordera la méthode de Chain Ladder. Pour l'instant nous allons nous contenter de la définition directe.

Pour revenir à la cadence cumulée PC_j relative au délai j , nous écrivons : $PC_j = \frac{C_{ij}}{C_{in}}$. Nous pouvons également trouver un lien entre les cadences de règlement et les facteurs de développement. En effet : $PC_j = \frac{1}{(F_j * F_{j+1} * \dots * F_{n-1})}$ pour $j= n-i \dots n-1$.

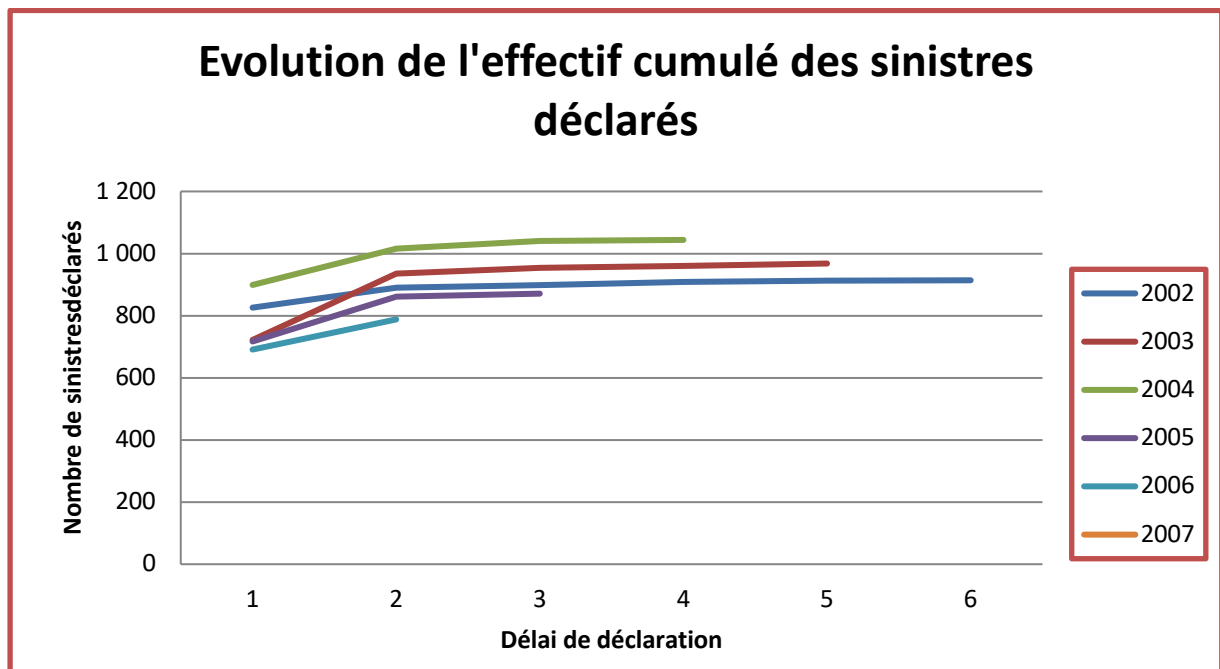
PC_j est de ce fait, pour toutes années d'origine confondues, le pourcentage cumulé des sinistres payés. Sinon, si on ne considère qu'une seule année de survenance, on dirait que la cadence de règlement PC_j n'est autre que le pourcentage réglé de la sinistralité correspondant à une année i complètement déroulée.

j	0	1	2	3	4	5	6	7	8	9
PC _j	6,3%	29,8%	59,5%	77,5%	87,2%	91,9%	95,7%	97,9%	99,9%	100,0%

Suite aux résultats du tableau, nous sommes en mesure de confirmer les constats de l'analyse graphique. Comme nous l'avons cité auparavant, les trois premières années renferment la part la plus importante de l'ensemble des paiements. En effet, 60% des règlements à effectuer sont faits durant les trois premières années de déroulement. Il est par ailleurs constaté qu'un retard de règlement s'impose pour la première année où juste 6,3% des sinistres sont en règle. Même la 2^{ème} année ne correspond qu'à 30% des règlements devant être effectués. Cela laisse penser à une lenteur dans la procédure d'indemnisation dont il faut examiner les causes.

II.3 Etude de la déclaration des sinistres

Nous en venons à l'analyse de la rapidité de la déclaration des sinistres. Pour cela, nous choisissons de faire la représentation graphique du nombre de sinistres ventilés selon l'année de survenance en fonction du délai pris pour faire la déclaration.



L'allure de toutes les courbes est similaire : elles ont toutes une tendance remarquable à la hausse pour les deux premières années avant de connaître une stagnation par la suite. On en déduit que la déclaration de la sinistralité relatives aux accidents de travail se fait principalement deux ans après la survenance ; on en conclut que les retards de paiements sont en partie dus aux déclarations non immédiates.

Conclusion

Après avoir explicité les principaux concepts et donné une plateforme permettant d'entamer la partie technique du projet, cette première section s'achève sur deux résultats majeurs :

- Les sinistres ne se règlent pas automatiquement après leur survenance. Effectivement, 91% des victimes d'un AT sont indemnisées 5 ans après qu'elles aient subies le préjudice en question. Incontestablement, la déclaration d'un sinistre se fait généralement au cours des deux premières années, ce qui fait que le processus de remboursement prend à lui seul 3 ans supplémentaires. Ainsi, le retard constaté est le résultat d'une déclaration tardive et d'une procédure d'indemnisation un peu lente.
- L'inversion du cycle de production et la longévité des engagements fait que la mise au point de la stratégie prudentielle de la compagnie repose sur des estimations et prévisions économiques, notamment lorsqu'il s'agit de constituer ses réserves. Il se trouve que les plus importantes réserves sont les provisions techniques ; elles renvoient directement à la crédibilité de l'assureur et à sa capacité d'honorer ses engagements et promesses envers ses assurés. En particulier, la prédominance volumique des PSAP suscite une attention singulière.

Partie II. Calculs non stochastiques des PSAP

Cette partie traite une multitude de méthodes déterministes et d'inventaires pour le calcul de provisions. Elle se subdivise en deux chapitres.

Le premier chapitre explicite les diverses étapes du calcul des PSAP dictées par le schéma réglementaire marocain. Il présente par ailleurs le résultat final de ce calcul.

Le second chapitre quant à lui regroupe un ensemble de méthodes non stochastiques, explicite leur principe de base et présente le résultat de chacune de ces méthodes. Ce chapitre établit également une étude comparative des résultats obtenus.

Introduction

Les pouvoirs publics ont édicté un ensemble de dispositions comptables et de règles de gestion prudentielles et préventives pour assurer la pérennité des compagnies d'assurances. Effectivement, la DAPS, première entité chargée du contrôle de l'activité des organismes d'assurances et de prévoyance sociale, participe à la codification des moindres détails relatifs au placement des fonds recueillis par ces organismes et au contrôle de leur gestion technique et financière. La DAPS veille spécialement au respect des droits des assurés et bénéficiaires de contrats d'assurances.

Dans ce cadre, que la réglementation a prévu une suite de méthodes permettant d'avoir une estimation des PSAP. Le schéma préconisé pour cette estimation est dûment expliqué et ne laisse point place à l'ambiguïté ; il fournit une procédure claire visant à tracer le processus d'évaluation des réserves techniques. Cette procédure n'implique en rien l'usage des calculs de probabilités et se base uniquement sur des algorithmes arithmétiques simples.

Les provisions pour sinistres à payer peuvent également être évaluées à l'aide d'une multitude de méthodes déterministes autres que celles exigées par l'autorité de tutelle. Ces méthodes, malgré leur diversité, se partagent un certain nombre de caractères relatifs à leur mise en œuvre, aux hypothèses fondamentales, l'idée basique derrière la construction de l'algorithme correspondant... Leur similitude la plus marquante est leur incapacité à donner une mesure du risque pour l'estimation faite. De ce fait, elles ne fournissent que les valeurs ponctuelles des réserves annuelles et de la réserve globale à constituer au titre de l'exercice considéré.

Dans une optique organisationnelle, il a été convenu de répartir les méthodes déterministes de provisionnement en plusieurs sous-groupes. Chaque ensemble regroupe des techniques très similaires. Ainsi, toutes les variantes de la méthode de Chain Ladder se trouvent regroupées, les méthodes des ratios sont également un sous-ensemble à part et reposent dans leurs algorithmes sur des ratios spécifiques, les méthodes autorégressives qui s'articulent autour de formules de calculs purement autorégressives...

Malgré cette multiplicité de techniques, on peut considérer que si plusieurs méthodes donnent des résultats convergents, ces résultats donnent une bonne approximation du montant des engagements.

Chapitre I. Le provisionnement selon les méthodes réglementaires

I. Le principe de calcul

L'évaluation des réserves techniques en assurance non-vie s'effectue selon l'exigence réglementaire marocaine en se basant sur la méthode dossier par dossier. Cette dernière définit la PSAP comme étant la somme des trois termes suivants :

- L'évaluation de base pour la provision à payer est dégagée dossier par dossier. C'est une technique d'inventaire intégrant bon nombre de concepts de comptabilité analytique et fait appel aux compétences de l'expert sinistre. Ce dernier doit intégrer dans son évaluation, en plus de la charge objective figurant dans la documentation, sa propre estimation de tout facteur pouvant accroître le coût du sinistre évalué. Il s'agit d'une lecture directe de la base des données :

Evaluations D/D	
Exercices antérieurs (1977-1997)	3919
1998	1252
1999	952
2000	963
2001	1124
2002	2955
2003	9765
2004	6794
2005	9447
2006	12317
2007	20480

- Le montant estimé pour la réserve des tardifs IBNR prend en charge les sinistres survenus mais non déclarés à la date de clôture de l'exercice.
- La réserve complémentaire est une mesure corrective permettant de pallier à une insuffisance de provision.

II. L'estimation de la réserve des tardifs

La valeur estimée de la réserve des tardifs est égale au produit du nombre des sinistres non déclarés et du coût moyen de ces derniers. Donc, pour calculer cette provision, on sera amené à déterminer le nombre des sinistres survenus non déclarés par exercice de survenance ainsi que le coût moyen des sinistres.

II.1 Le calcul du coût moyen

Pour estimer le coût moyen, on procède par la division de la charge totale des sinistres terminés au cours des cinq dernières années par leur effectif :

$$\text{Coût moyen} = \frac{\text{le coût des dossiers terminés au cours des 5 dernières années}}{\text{le nombre de sinistres terminés ou classés sans suite au cours de la même période}}$$

Par application directe de la formule ci-dessus sur les données dont nous disposons, on trouve le résultat suivant :

	2007	2006	2005	2004	2003	TOTAL
Coût des sinistres terminés	10409	15577	12365	9219	8387	55957
Nombre des sinistres terminés	1127	1495	1112	734	863	5331

N.B : Le coût des sinistres terminés est exprimé en milliers de dirhams.

II.2 L'estimation de l'effectif des tardifs et la conclusion de la réserve IBNR

L'estimation du nombre des tardifs se fait par la technique de Chain Ladder. Mais avant tout développement de ladite méthode, il semble adéquat d'introduire la forme générale du tableau triangulaire donnant le nombre des sinistres déclarés par année de survenance et par année de déclaration.

		Délai de déclaration (j)						
		0	1	2	3		n-1	n
Année de survenance(i)	0	N ₀₀	N ₀₁	N ₀₂	N ₀₃		N _{0n-1}	N _{0n}
	1	N ₁₀	N ₁₁	N ₁₂	N ₁₃		N _{1n-1}	
	2	N ₂₀	N ₂₁	N ₂₂	N ₂₃			
	3	N ₃₀	N ₃₁	N ₃₂	N ₃₃			
	n-1	N _{n-1,0}	N _{n-1,1}					
	n	N _{n0}						

Pour tracer le tableau des effectifs cumulés associé au triangle de données ci-dessus, on utilise les formules suivantes :

$$\begin{cases} C_{i0} = N_{i0} \\ C_{ij} = N_{ij} + C_{i,j-1} \quad j > 0 \end{cases}$$

Par la suite, pour compléter le tableau cumulé, on calcule les facteurs d'aggravation notés F_j. Ces facteurs s'interprètent comme étant des coefficients de passage d'une colonne à une autre. On rappelle que le facteur correspondant au passage de la colonne j à la colonne j+1 s'écrit :

$$F_j = \frac{\sum_{i=0}^{i=n-j-1} C_{i,j+1}}{\sum_{i=0}^{i=n-j-1} C_{ij}}$$

En effet, pour compléter la j^{ème} colonne on multiplie la (j-1)^{ème} colonne par le facteur de développement F_{j-1}. Du reste, l'estimation de nombre des tardifs pour chaque année de survenance se fait par un simple calcul : le nombre des tardifs pour une année i et après k années de déroulement est le cumul C_{in} diminué de C_{ik}.

Les facteurs de développement calculés sont :

J	0	1	2	3	4
F _j	1,1650	1,0170	1,0066	1,0053	1,0022

Ces coefficients de passage nous permettent de compléter le tableau mais surtout d'avoir la dernière colonne :

EXERCICES DE DECLARATION EXERCICES DE SURVENANCE	0	1	2	3	4	5	Valeur ultime	Elément de la diagonale
2002	826	890	899	909	912	914	914	914
2003	722	935	955	961	968		970	968
2004	899	1 016	1 041	1 044			1 052	1 044

2005	717	862	871	883	871
2006	691	788		813	788
2007	836			1 005	836

Suite à ce tableau on déduira le nombre des tardifs par année de survenance : c'est la différence entre la valeur ultime et l'élément de la diagonale.

Pour calculer la réserve des tardifs on multiplie leur effectif par le coût moyen déjà estimé :

EXERCICES DE SURVENANCE	tardifs	CM	réserve
2002	0	10,5	0,00
2003	2	10,50	20,99
2004	8	10,50	83,97
2005	12	10,50	125,96
2006	25	10,50	262,41
2007	169	10,50	1773,91
		TOTAL	2267,25

Remarque : On s'est limité au calcul de la réserve des tardifs pour les 5 années qui précèdent l'année 2007 pour un manque dans les données et on suppose que le délai de déclaration d'un sinistre AT est de 5 ans. Ainsi, toutes les années précédant l'année 2002 ont une provision IBNR nulle.

III. Le calcul de la réserve complémentaire

Après avoir estimé la réserve des tardifs, on s'intéresse à l'estimation de la réserve complémentaire. Cette dernière peut s'assimiler à un seuil minimal garantissant plus de sécurité quant à l'estimation de la réserve totale.

L'évaluation de la réserve complémentaire requiert, toujours selon les préconisations réglementaires, l'usage des méthodes qu'on expliquera ci-après ; à savoir la méthode du coût moyen et celle de la cadence de règlement. Par la suite, si le résultat obtenu par l'une ou l'autre des deux méthodes est plus important que l'évaluation calculée auparavant, cette dernière est complétée de la différence entre elle et le résultat le plus élevé.

III.1 La méthode du coût moyen

Cette méthode consiste à multiplier le nombre total des sinistres survenus (y compris l'estimation de ceux non déclarés à la date de l'inventaire) par le coût moyen des sinistres déjà calculé.

La réglementation en vigueur définit les montants suivants :

- La réserve résiduelle = La réserve D/D + La réserve des tardifs
- La charge des sinistres = La réserve résiduelle + Le cumul des règlements

Dans le cas où la provision résiduelle est inférieure ou égale à 30% de la charge des sinistres, la valeur calculée est remplacée par la somme de l'évaluation D/D et la réserve des tardifs. Cela dit, cette méthode n'est applicable que pour dix exercices au plus.

	nb.sin survenus	CM	nb * CM	Cumul des règlements	Provision CM	réserve résiduelle	charge sinistres	RES/CHG	Provision
1998	960	10,50	10076,67	14643	-4566,33	1252	15895	-	1252
1999	946	10,50	9929,72	10669	-739,28	952	11621	-	952
2000	1028	10,50	10790,43	11902	-1111,57	963	12865	-	963
2001	871	10,50	9142,48	9132	10,48	1124	10256	10,96%	1124

2002	914	10,50	9593,83	9845	-251,17	2955	12800	-	2955
2003	970	10,50	10181,63	16614	-6432,37	9786	26400	-	9786
2004	1052	10,50	11042,35	15930	-4887,65	6878	22808	-	6878
2005	883	10,50	9268,44	9236	32,44	9573	18809	50,90%	32,44
2006	813	10,50	8533,68	3758	4775,68	12579,5	16337	77,00%	4775,68
2007	1005	10,50	10549,01	1122	9427,01	22254,5	23376	95,20%	9427,01

III.2 La méthode des cadences de règlements

Cette méthode nécessite d'avoir un tableau qui regroupe le cumul des sinistres payés ainsi que la charge des sinistres par année de survenance sur une période d'au moins 10 ans. En se basant sur les données de ce tableau, on calcule la cadence et par la suite on en déduit la réserve. Les formules utilisées sont :

$$\checkmark \text{ Cadence} = \frac{\text{réglements cumulés}}{\text{charges de sinistres}}$$

$$\checkmark \text{ la réserve} = \text{réglements cumulés} * \left(\frac{100 - \text{cadence}}{\text{cadence}} \right)$$

Remarque : La formule donnant la réserve se base sur l'idée que le ratio des pourcentages est égal à celui des montants et on peut la réécrire comme suit :

$$\frac{\text{La réserve totale (ce qui reste à payer)}}{\text{Réglements cumulés (ce qui a été payé)}} = \frac{100 - \text{cadence (le pourcentage des non payés)}}{\text{cadence (Le pourcentage des payés)}}$$

	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998
Cadence	5,27%	25,63%	51,13%	66,78%	74,71%	83,69%	88,71%	90,90%	91,95%	92,12%
100%-cadence	94,73%	74,37%	48,87%	33,22%	25,29%	16,31%	11,29%	9,10%	8,05%	7,88%
cumul règlement	1 122	3 758	9 236	15 930	16 614	9 845	9 132	11 902	10 669	14 643
Réserve	20151	10905	8828	7925	5625	1919	1162	1191	934	1252

Réserve totale	59 892
----------------	--------

III.3 Calcul de la réserve complémentaire et déduction de la réserve finale

La constitution de la réserve complémentaire a un objectif ultime : s'assurer de la suffisance des fonds provisionnés. Cela fait que le coût moyen et celle la cadence de règlements ne sont pas des méthodes à part entière ; elles ne viennent que donner plus de crédibilité aux résultats obtenus par la méthode de base, à savoir la méthode D/D.

Après avoir appliqué la méthode du coût moyen et la méthode de la cadence de règlement, on calcule la réserve complémentaire. En effet :

$\text{Provision complémentaire} = \text{Max} (\text{Max} (\text{réserve coût moyen, réserve cadence de règlements}) - \text{D/D} + \text{tardifs}, 0)$

La suite est un rapprochement des trois méthodes et la déduction de la réserve PSAP à retenir. Les résultats obtenus à ce niveau sont regroupés dans le tableau ci-dessous :

	D/D + tardifs	Coût moyen	Cadence	Réserve complémentaire
1 997	3 919	3 919	3 919	
1 998	1 252	1 252	1 252	
1 999	952	952	934	
2 000	963	963	1 191	
2 001	1 124	1 124	1 162	

2 002	2 955	2 955	1 919	
2 003	9 786	9 786	5 625	
2 004	6 878	6 878	7 925	
2 005	9 573	32	8 828	
2 006	12 580	4 776	10 905	
2 007	22 255	9 427	20 151	
Total	72 237	42 064	63 811	0,00

PSAP réglementaire	72 237
---------------------------	---------------

L'exposé théorique et numérique du schéma réglementaire marocain en matière de constitution des PSAP nous a menées à une estimation de la réserve de 72 237.

Tout au long de l'application des méthodes imposées par l'autorité de tutelle, on constate que ces techniques sont carrément comptables et visent avant tout à avoir une provision conforme aux exigences de réglementation. Cependant, il est à noter qu'une certaine marge d'appréciation subjective est laissée à l'assureur grâce à la méthode d'inventaire D/D. Ainsi, le calcul des PSAP donne lieu à une grande marge d'erreur en impliquant l'expérience personnelle du gestionnaire de risque et son habilité à évaluer avec acuité les risques futurs.

Chapitre II. Les méthodes déterministes du calcul des réserves

I. Les méthodes de Chain Ladder

Les méthodes Chain Ladder font partie des méthodes déterministes les plus connues et les plus répandues dans le calcul des provisions. Particulièrement simples et directes, ces méthodes ne requièrent qu'une base de données triangulaire, éventuellement le triangle de liquidation. L'idée de base repose sur l'usage des facteurs de développement déjà définis et ce, dans le but de prédire les données futures manquantes rien qu'en se référant à celles déjà observées.

Pour chacune des variantes de Chain Ladder, des hypothèses de départ doivent être supposées. Cependant, après avoir abouti aux résultats voulus, il est nécessaire de revoir la justesse de nos suppositions, surtout lorsqu'il s'agit d'hypothèses fortes.

Tout au long de cette section, on reste fidèle aux notations suivantes :

- I est l'indice de l'année de survenance.
- J est l'indice du délai de déroulement.
- C_{ij} est le le montant des paiements cumulé jusqu'en l'année de développement j des sinistres survenus en l'année d'accident i.
- N est l'indice de l'exercice dont on cherche la provision. Puisque nos indices commencent avec un zéro, nous travaillons donc sur n+1 années.

I.1 La méthode de Chain Ladder Standard :

Cette méthode est la plus fréquente pour le calcul des PSAP. Quoiqu'elle repose sur un algorithme simpliste, la compréhension des enjeux de sa démarche demeure cruciale.

I.1.1 Les hypothèses de base

La méthode de Chain Ladder repose sur une première hypothèse d'indépendance des années de survenance des sinistres. En pratique, cette est généralement admise et ne nécessite pas une vérification précise.

Par ailleurs, les facteurs de développement étant grandement impliqués dans la procédure utilisée par Chain Ladder Standard, ils font l'objet d'une forte hypothèse indispensable à la logique de cette méthode. En effet, implicitement, la méthode Chain Ladder Standard suppose que les ratios $\frac{C_{i,j+1}}{C_{i,j}}$ sont indépendants de l'année d'origine i. Autrement, cela signifie que le pourcentage des sinistres liquidés entre deux années comptables successives ne dépend pas de l'année de survenance de ces sinistres. Ainsi, pour tout $i=0, \dots, n-j-1$ nous avons $\frac{C_{i,j+1}}{C_{i,j}} = \text{cte}$ qu'on note F_j . Cette constatation fera certainement l'objet de la vérification de l'hypothèse d'indépendance.

I.1.2 L'estimation de la réserve

Il s'agit de calculer les éléments manquants du tableau en estimant les paiements futurs pour aboutir à un rectangle complet de liquidation.

En pratique, l'hypothèse d'indépendance n'étant qu'approximativement vérifiée, nous avons recours à des facteurs plus synthétiques et communs. Il s'agit des facteurs définis tels que :

$$F_j = \frac{\sum_{i=0}^{i=n-j-1} C_{i,j+1}}{\sum_{i=0}^{i=n-j-1} C_{i,j}}$$

A l'aide de ces coefficients de passage communs, on procède à l'estimation des montants manquants comme suit :

$$C_{ij} = C_{i,n-i} * f_{n-i} * \dots * f_{j-1} = C_{i,n-i} \prod_{h=n-i}^{j-1} f_h \text{ pour le cas où on a } i+j > n$$

Notons que la charge ultime C_{in} est égale à la charge totale S_i de tous les sinistres survenus dans la même année i .

L'étape suivante consiste à retrouver la provision globale pour l'année considérée. Elle est égale à la somme des réserves calculées par année de survenance. Si on note R la provision globale et R_i la provision relative à l'année i on aura :

$$\begin{cases} R_i = C_{in} - C_{i,n-i} \\ R = \sum_{i=1}^n R_i \end{cases}$$

Les coefficients de développement se présentent ainsi :

j	0	1	2	3	4	5	6	7	8	9
F_j	4,7133	1,9971	1,3011	1,1260	1,0535	1,0419	1,0225	1,0203	1,0009	1,0003

Suite à une application directe des formules précédentes, on complète les éléments manquants du tableau pour aboutir aux valeurs des réserves annuelles et à l'estimation de la valeur globale :

	0	1	2	3	4	5	6	7	8	9	10	R_i
1997	792	3 119	7 155	9 522	10 370	11 036	11 631	11 783	12 029	12 041	12 045	-
1998	847	4 380	8 977	10 908	12 370	12 831	13 392	14 143	14 631	14 643	14 647,86	4,86
1999	708	3 253	6 300	8 179	9 718	10 287	10 518	10 662	10 669	10 678,60	10 682,15	13,15
2000	933	3 378	6 683	9 282	10 753	11 473	11 880	11 902	12 143,05	12 153,98	12 158,01	256,01
2001	783	3 291	5 852	7 393	8 202	8 653	9 132	9 337,86	9 526,98	9 535,55	9 538,72	406,72
2002	783	3 429	6 557	8 159	9 458	9 845	10 257,26	10 488,49	10 700,91	10 710,54	10 714,10	869,10
2003	816	6 227	2 074	15 372	16 614	17 502,14	18 235,05	18 646,12	19 024	19 040,88	19 047,20	2 433,20
2004	1 156	5 154	11 534	15 930	17 937,02	18 895,89	19 687,16	20 130,96	20 538,66	20 557,15	20 563,98	4 633,98
2005	903	5 007	9 236	12 017,21	13 531,26	14 254,60	14 851,52	15 186,31	15 493,87	15 507,82	15 512,97	6 276,97
2006	977	3 758	7 505	9 765,09	10 995,39	11 583,18	12 068,23	12 340,28	12 590,20	12 601,53	12 605,72	8 847,72
2007	1 122	5 288	10 561	13 741,51	15 472,80	16 299,93	16 982,50	17 365,33	17 717,02	17 732,97	17 738,87	16 616,87
												40 358,60

Ainsi la réserve retenue d'après la méthode de Chain Ladder Standard est 40 358,60

1.1.3 La validation de l'hypothèse d'indépendance des coefficients de passage

Il existe deux façons pour vérifier l'exactitude de l'hypothèse d'indépendance des facteurs de développement. La première est purement graphique alors que la deuxième repose plutôt sur le calcul et l'interprétation de quelques statistiques de base.

1) La vérification par C-C Plot

Suivant cette méthode, on se doit de générer, pour chaque j fixé, un graphe de couples $(C_{ij}, C_{i,j+1})$, i allant de 0 jusqu'à $n-j-1$. Pour valider l'hypothèse, il suffit de trouver un nuage de points s'alignant autour d'une droite fictive passant pas l'origine.

En effet, on constate sur les diagrammes C-C Plot générés que les nuages de points décrivent la silhouette d'une droite passant par l'origine (voir annexe 1). De ce fait, on confirme la justesse de notre hypothèse.

2) La vérification par l'usage de statistiques

Une autre alternative pour tester l'hypothèse d'indépendance repose sur l'usage du D-triangle. Ce dernier n'est autre qu'un tableau triangulaire formé des coefficients de passage déjà définis. Pour être plus de précis, notre objectif est de calculer sur chaque colonne du D-tableau que l'on va définir, quelques statistiques basiques, à savoir l'espérance (m), l'écart-type (σ) et le coefficient de variation (cv).

L'idéal serait de pouvoir affirmer avoir une volatilité très faible grâce à un cv décroissant et insignifiant.

0	1	2	3	4	5	6	7	8	9
3,9381	2,2940	1,3308	1,0891	1,0642	1,0539	1,0131	1,0209	1,0010	1,0003
5,1712	2,0495	1,2151	1,1340	1,0373	1,0437	1,0561	1,0345	1,0008	
4,5946	1,9367	1,2983	1,1882	1,0586	1,0225	1,0137	1,0007		
3,6206	1,9784	1,3889	1,1585	1,0670	1,0355	1,0019			
4,2031	1,7782	1,2633	1,1094	1,0550	1,0554				
4,3793	1,9122	1,2443	1,1592	1,0409					
7,6311	1,9390	1,2731	1,0808						
4,4585	2,2379	1,3811							
5,5449	1,8446								
3,8465									

Le tableau des statistiques sur lesquelles nous allons nous baser pour vérifier la véracité de l'hypothèse d'indépendance est le suivant :

	0	1	2	3	4	5	6	7	8	9
m	4,7388	1,9967	1,2994	1,1313	1,0538	1,0422	1,0212	1,0187	1,0009	1,0003
σ	1,1138	0,1615	0,0590	0,0370	0,0111	0,4256	0,4571	1,0188	0,0001	0,0000
cv	0,2350	0,0809	0,0454	0,0327	0,0106	0,4084	0,4476	1,0001	0,0001	0,0000

Le coefficient de corrélation est relativement petit pour les premières colonnes et il continue à diminuer jusqu'à atteindre la valeur 0. Il est donc clair que nous avons une volatilité négligeable et que par conséquent, l'indépendance des facteurs de développement est établie.

I.2 Les méthodes de Chain Ladder pondérées

Ici, nous sommes en présence d'une autre variante de la méthode de Chain Ladder. Le principe de calcul global est le même, seule diffère la procédure derrière le choix des facteurs de développement à retenir pour les calculs de la provision finale. Autrement dit, les méthodes de Chain Ladder pondérées gardent la même logique que Chain Ladder standard et suivent les mêmes étapes dans le calcul sauf que les facteurs de développement sont choisis d'une autre manière.

L'étude commence par le D. triangle défini et construit dans le paragraphe précédent. Sur une même colonne de ce triangle, des opérations sont faites sur les facteurs de développement pour aboutir à un facteur ultime qui sera représentatif du passage entre les deux délais consécutifs en question. Selon l'aspiration de chaque méthode de calcul, nous choisissons la combinaison des coefficients qui sied le mieux :

- La moyenne: on peut avoir recours à la moyenne si l'on veut inclure toutes les données même les plus anciennes. Le calcul de cette moyenne s'avère particulièrement judicieux quand la gestion de la sinistralité a été constante durant toutes les années considérées.

- La moyenne des trois derniers : on calcule ici la moyenne arithmétique des trois derniers facteurs de développement ; c'est-à-dire que, pour un même délai, on ne tient compte que des trois dernières années comptables disponibles. Le calcul en est ainsi plus abrégé mais la représentativité des données est moins présente, seules mises en valeur les données les plus actuelles.
- Le dernier facteur.

1.2.1 La moyenne des facteurs

Sur la base du D.triangle, on calcule le tableau des moyennes par année de développement :

j	0	1	2	3	4	5	6	7	8	9
m	4,7388	1,9967	1,2994	1,1313	1,0538	1,0422	1,0212	1,0187	1,0009	1,0003

Le triangle ainsi complété se présente comme suit :

	0	1	2	3	4	5	6	7	8	9	10	provision
1997	792	3 119	7 155	9 522	10 370	11 036	11 631	11 783	12 029	12 041	12 045	-
1998	847	4 380	8 977	10 908	12 370	12 831	13 392	14 143	14 631	14 643	14 647,9	4,9
1999	708	3 253	6 300	8 179	9 718	10 287	10 518	10 662	10 669	10 678,7	10 682,2	13,2
2000	933	3 378	6 683	9 282	10 753	11 473	11 880	11 902	12 124,3	12 135,3	12 139,4	237,4
2001	783	3 291	5 852	7 393	8 202	8 653	9 132	9 325,4	9 499,5	9 508,2	9 511,3	379,3
2002	783	3 429	6 557	8 159	9 458	9 845	10 260,3	10 477,6	10 673,3	10 683	10 686,5	841,5
2003	816	6 227	12 074	15 372	16 614	17 508,1	18 246,7	18 633	18 981	18 998,3	19 004,6	2 390,6
2004	1 156	5 154	11 534	15 930	18 021,8	18 991,6	19 792,8	20 211,9	20 589,4	20 608,1	20 615	4 685
2005	903	5 007	9 236	12 001,1	13 576,9	14 307,6	14 911,1	15 226,8	15 511,3	15 525,4	15 530,5	6 294,5
2006	977	3 758	7 504	9 750,1	11 030,4	11 624	12 114,4	12 370,8	12 601,9	12 613,4	12 617,6	8 859,6
2007	1 122	5 317	10 616	13 794,7	15 606,1	16 445,9	17 139,7	17 502,6	17 829,5	17 845,7	17 851,7	16 729,7
												40435,7

La réserve pour sinistres à payer donnée par cette variante de Chain Ladder est de 40 435,7.

1.2.2 La moyenne des trois derniers facteurs

On procède de la même manière que précédemment pour calculer en premier lieu les facteurs de développement uniquement sur la base des trois dernières observations.

j	0	1	2	3	4	5	6	7	8	9
M ₃	4,6166	2,0072	1,2995	1,1165	1,0543	1,0378	1,0239	1,0187	1,0009	1,0003

Le tableau complété et les réserves calculées suite à ces facteurs se présentent comme suit :

	0	1	2	3	4	5	6	7	8	9	10	provision
1997	792	3 119	7 155	9 522	10 370	11 036	11 631	11 783	12 029	12 041	12 045	-

1998	847	4 380	8 977	10 908	12 370	12 831	13 392	14 143	14 631	14 643	14 647,9	4,9
1999	708	3 253	6 300	8 179	9 718	10 287	10 518	10 662	10 669	10 678,7	10 682,2	13,2
2000	933	3 378	6 683	9 282	10 753	11 473	11 880	11 902	12 124,3	12 135,3	12 139,4	237,4
2001	783	3 291	5 852	7 393	8 202	8 653	9 132	9 350	9 524,7	9 533,3	9 536,5	404,5
2002	783	3 429	6 557	8 159	9 458	9 845	10 216,8	10 460,7	10 656,1	10 665,8	10 669,3	824,3
2003	816	6 227	12 074	15 372	16 614	17 515,9	18 177,4	18 611,3	18 959	18 976,2	18 982,5	2 368,5
2004	1 156	5 154	11 534	15 930	17 785,5	18 751	19 459,1	19 923,7	20 295,8	20 314,3	20 321	4 391
2005	903	5 007	9 236	12 002,5	13 400,5	14 128	14 661,5	15 011,5	15 291,9	15 305,8	15 310,9	6 074,9
2006	977	3 758	7 543	9 802,2	10 944	11 538,1	11 973,8	12 259,7	12 488,7	12 500	12 504,2	8 746,2
2007	1 122	5 180	10 397	13 510,9	15 084,6	15 903,5	16 504,1	16 898,1	17 213,7	17 229,4	17 235,1	16 113,1
39178,1												

La réserve pour sinistres à payer donnée par cette autre variante de Chain Ladder est de 39 178,1.

1.2.3 Le dernier facteur

Le tableau suivant regroupe les facteurs de développement au titre de la dernière année comptable et ce, pour chaque délai de déroulement:

j	0	1	2	3	4	5	6	7	8	9
F _j	3,8465	1,8446	1,3811	1,0808	1,0409	1,0554	1,0019	1,0007	1,0008	1,0003

Le triangle complété et la réserve ultime sont résumés dans le tableau suivant:

	0	1	2	3	4	5	6	7	8	9	10	provision
1997	792	3 119	7 155	9 522	10 370	11 036	11 631	11 783	12 029	12 041	12 045	-
1998	847	4 380	8 977	10 908	12 370	12 831	13 392	14 143	14 631	14 643	14 647,9	4,9
1999	708	3 253	6 300	8 179	9 718	10 287	10 518	10 662	10 669	10 677,8	10 681,3	12,3
2000	933	3 378	6 683	9 282	10 753	11 473	11 880	11 902	11 909,8	11 919,6	11 923,5	21,5
2001	783	3 291	5 852	7 393	8 202	8 653	9 132	9 148,9	9 154,9	9 162,4	9 165,5	33,5
2002	783	3 429	6 557	8 159	9 458	9 845	10 390	10 409,2	10 416,1	10 424,6	10 428,1	583,1
2003	816	6 227	12 074	15 372	16 614	17 293,8	18 251,2	18 284,9	18 297	18 311,9	18 318	1 704
2004	1 156	5 154	11 534	15 930	17 217,1	17 921,6	18 913,6	18 948,7	18 961,1	18 976,7	18 983	3 053
2005	903	5 007	9 236	12 756,2	13 786,8	14 350,9	15 145,4	15 173,4	15 183,4	15 195,8	15 201	5 965
2006	977	3 758	6 932	9 574,1	10 347,7	10 771,1	11 367,3	11 388,4	11 395,9	11 405,2	11 409	7 651
2007	1 122	4 316	7 961	10 995,1	11 883,4	12 369,7	13 054,4	13 078,6	13 087,2	13 097,9	13 102	11 980,2
31008,3												

La réserve pour sinistres à payer donnée par cette autre variante de Chain Ladder est de 31 008,3.

1.2.4 Comparaison des résultats des méthodes de Chain Ladder

A ce stade, une petite comparaison des résultats obtenus par les différentes méthodes s'impose. Effectivement, la table suivante est un récapitulatif des différentes provisions fournies selon chaque méthode.

Exercice	Chain-ladder standard	moyenne	moyenne. 3 derniers	Le dernier facteur
1997	-	-	-	-
1998	4,86	4,86	4,86	4,86
1999	13,15	13,24	13,24	12,30
2000	256,01	237,38	237,38	21,54
2001	406,72	379,33	404,49	33,47
2002	869,10	841,51	824,31	583,07
2003	2433,20	2390,64	2368,52	1704,03
2004	4633,98	4684,97	4391,03	3052,97
2005	6276,97	6294,50	6074,93	5964,86
2006	8847,72	8859,56	8746,18	7650,99
2007	16616,87	16729,66	16113,09	11980,24
	40358,60	40435,66	39178,04	31008,32

Voici quelques quelques caractéristiques des ces quatre estimations :

provision moyenne	37745,15
écart-type	3921,32
cv	10,39%

II. Les méthodes autorégressives

Les méthodes de Chain Ladder supposent à l'unanimité que pour une année de déroulement fixée j dans $[0, n-1]$, le nuage des points $((C_{ij}, C_{i,j+1}))$, où i se trouve dans l'intervalle $[0, n-j-1]$ forme une droite passant par l'origine. Il s'avère en effet que cette hypothèse est assez forte et qu'elle ne peut être vérifiée sur tous les jeux de données.

Pour alléger cette supposition contraignante et pour élargir le nombre des bases de données sur lesquelles on peut travailler, les méthodes autorégressives supposent judicieusement que pour un j fixé, $C_{i,j+1}$ est une fonction affine de C_{ij} .

II.1 La méthode de London Chain

Ce modèle suppose, contrairement à Chain Ladder, que les points $(C_{i,j+1}, C_{i,j})$ sont sensiblement alignés autour d'une droite qui ne passe pas obligatoirement par l'origine. Cela signifie que la dynamique des $C_{i,j+1}$ est de la forme suivante : $C_{i,j+1} = f_j * C_{i,j} + a_j$ appartenant à $[0, n-j-1]$.

Cette relation s'assimile à une relation de régression linéaire simple où :

- f_j : est le paramètre de la régression linéaire simple.
- a_j : est la constante de régression.

Donc pour estimer le couple (f_j, a_j) on procède par la méthode des Moindres Carrés Ordinaires MCO. Dans ce sens, on doit minimiser la fonction suivante :

$$D_j = \sum_{i=0}^{n-j-1} (C_{i,j+1} - f_j * C_{i,j} - a_j)^2$$

On obtient alors les estimateurs classiques de la régression simple ci-dessous :

- $f_j = \frac{\frac{1}{n-j} \sum_{i=0}^{n-j-1} (C_{ij} * C_{i,j+1} - \bar{C}_j * \bar{C}_{j+1})}{\frac{1}{n-j} \sum_{i=0}^{n-j-1} (C_{ij})^2 - \bar{C}_j^2}$
- $a_j = \bar{C}_{j+1} - f_j * \bar{C}_j$

Avec : $\bar{C}_j = \frac{1}{n-j} \sum_{i=0}^{n-j-1} C_{ij}$ et $\bar{C}_{j+1} = \frac{1}{n-j} \sum_{i=0}^{n-j-1} C_{i,j+1}$

En outre, on a convenu du fait que le dernier couple (f_{n-1}, a_{n-1}) ne respecte pas les formules de calcul ci-dessus. Donc, par convention $a_{n-1}=0$ et $f_{n-1} = \frac{\bar{C}_n}{\bar{C}_{n-1}}$

Le tableau suivant est un récapitulatif des des valeurs des couples (f_j, a_j) pour chaque j fixé :

j	0	1	2	3	4	5	6	7	8	9
f_j	3,1531	1,9963	1,3303	1,0161	1,0232	1,0295	1,2209	1,1327	1,0000	1,0003
a_j	1357,02	3,46	-237,33	1080,60	306,66	134,87	-2351,72	-1371,14	12,00	0,00

Après avoir calculé les couples (f_j, a_j) , on complète le triangle de paiement en se basant sur l'équation du modèle déjà motionnée et on en déduit la provision globale.

	0	1	2	3	4	5	6	7	8	9	10	provision
1997	792	3 119	7 155	9 522	10 370	11 036	11 631	11 783	12 029	12 041	12 045	-
1998	847	4 380	8 977	10 908	12 370	12 831	13 392	14 143	14 631	14 643	14 647,9	4,9
1999	708	3 253	6 300	8 179	9 718	10 287	10 518	10 662	10 669	10 681	10 684,6	15,6
2000	933	3 378	6 683	9 282	10 753	11 473	11 880	11 902	12 110	12 122	12 126	224
2001	783	3 291	5 852	7 393	8 202	8 653	9 132	8 797,7	8 593,8	8 605,8	49 799,6	40 667,6
2002	783	3 429	6 557	8 159	9 458	9 845	10 269,8	10 186,8	10 167,3	10 179,3	10 183,6	338,6
2003	816	6 227	12 074	15 372	16 614	17 306,6	17 951,2	19 565,1	20 789,8	20 801,8	20 808,7	4 194,7
2004	1 156	5 154	11 534	15 930	17 266,6	17 974,4	18 638,6	20 404,4	21 740,4	21 752,4	21 759,7	5 829,7
2005	903	5 007	9 236	12 049,1	13 323,3	13 939,5	14 484,9	15 333,1	15 996,3	16 008,3	16 013,6	6 777,6
2006	977	3 758	4 762	6 097,3	7 275,8	7 751,5	8 114,7	7 555,6	7 186,9	7 198,9	7 201,3	3 443,3
2007	1 122	2 243	2 747	3 416,7	4 552,2	4 964,6	5 245,7	4 052,9	3 219,4	3 231,4	3 232,5	2 110,5
												63606,5

Remarque: Les valeurs de la 7^{ème} colonne sont inférieures aux valeurs de la 6^{ème} colonne ce qui ne respecte pas le principe du triangle cumulé qui suppose que les valeurs de chaque ligne augmentent en passant d'une colonne à une autre.

II.2 La méthode de London Pivot :

Pour plus de commodité, nous avons inversé les notations établies pour le reste de la partie. Ainsi, i se trouve devenir l'indice du délai de développement et j celui de l'année d'origine. Le tableau qu'on va utiliser pour cette méthode sera ainsi l'inverse du triangle de liquidation qu'on a déjà étudié.

Cette technique suppose l'existence d'un paramètre « a » indépendant de j qui vérifie pour chaque valeur i de l'intervalle [0,n-1] l'équation suivante: $C_{j,i+1}+a=f_i(C_{ji}+a)$ pour $j=0.....n-1-i$. Donc pour mettre en pratique cette méthode on doit estimer les paramètres f_j ainsi que le paramètre a.

Pour réaliser nos objectifs, notamment celui de l'estimation de la réserve, on applique un algorithme spécifique qu'on peut résumer en quatre étapes essentielles :

Etape 1 : On donne une valeur initiale a_0 au paramètre a.

Etape 2 : On calcule $b_i(a) = \frac{\sum_{j=i+1}^n (C_{i+1,j}+a)(C_{ij}+a)}{\sum_{j=i+1}^n (C_{i+1,j}+a)^2}$

Etape 3 : On déduit la valeur suivante de a grâce à la formule:

$$a_1 = \frac{\sum_{i=1}^{n-1} (b_i - 1) (\sum_{j=i+1}^n C_{i+1,j} - b_i \sum_{j=i+1}^n C_{ij})}{\sum_{i=1}^n (n-i)(b_i - 1)^2}$$

Etape 4 : On répète ce algorithme depuis l'étape 1 en affectant à a_0 la valeur de a_1 . On continue à faire des boucles jusqu'à atteindre la valeur du paramètre « a » qui minimise l'expression suivante : $\sum_{i=1}^{n-1} \sum_{j=i+1}^n (C_{i+1,j} + a - b_i(C_{i,j} + a))^2$

Voici donc une table donnant quelques valeurs de a avec les valeurs de la provision correspondante :

La valeur de « a »	-10119,888	-10120,888	-10119
La provision	2260,663444	2259,854632	2261,374952

Vu la complexité de la mise en œuvre de cette méthode, il a été difficile de se décider quant à la valeur finale qui satisfait la condition de minimisation. Cependant, nous avons jugé que la valeur $a = -6,79142E+11$ est assez satisfaisante et les réserves qui en découlent sont :

Année de survenance	Ri
1 997	-
1 998	4,00
1 999	16,00
2 000	263,00
2 001	530,25
2 002	984,85
2 003	1 527,18
2 004	4 939,61
2 005	4 939,61
2 006	9 065,17
2 007	12 294,97
Total	34 564,64

La réserve pour sinistres à payer donnée par cette méthode est de 34 564,64.

III. Les méthodes basées sur les ratios :

Les méthodes basées sur des ratios utilisent en plus du triangle de liquidation, un supplément d'informations. En effet, la donnée d'un indicateur d'exposition au risque ou de sinistralité est indispensable pour l'application de telles méthodes. Dans le cas général, cet indicateur peut être le montant de primes, le nombre de sinistres, le nombre de contrats...

Dans notre étude, l'indicateur utilisé est le montant des primes acquises noté P_i . Par la suite, on adoptera les notations suivantes pour des ratios qu'on utilisera par la suite :

- $L_{ij} = \frac{C_{ij}}{P_i} \quad i, j = 0 \dots n$
- $L_i = \frac{C_{in}}{P_i} = \frac{S_i}{P_i} \quad i = 0 \dots n$ pour le ratio ultime.

Particulièrement, et puisque notre indicateur d'exposition au risque est une prime, le ratio $L_{i,n-i}$ n'est autre que la part de la prime acquise en l'année i consommée à la date d'inventaire de cette même année. Il s'agit exactement du ratio Sinistres à Primes de l'année i (dit encore Loss Ratio) qui est très utilisé en assurance non-vie.

III.1 La méthode du Loss Ratio simple

Cette méthode suppose que les Loss Ratios ultimes L_i sont constants. Elle tient sa spécificité du fait qu'elle applique un unique ratio quelque soit l'année de survenance et on

écrit : $L_i=L_0$ pour tout $i=0\dots n$. Cependant, cette technique n'est applicable que pour les branches totalement stables et il se peut qu'elle donne des résultats aberrants.

Pour calculer la provision selon cette méthode on commence par le calcul du ratio commun à toutes les années de survenance L_0 . Puis on en déduit la provision selon les formules suivantes :

- $S_i= P_i*L_0$
- $R_i=S_i - C_{i,n-i}$
- $R=\sum_{i=0}^n R_i$

	L₀	S_i	R_i
1997	93,7%	12 045	0
1998	93,7%	9 391	-5 252
1999	93,7%	13 242	2 573
2000	93,7%	11 320	-582
2001	93,7%	12 041	2 909
2002	93,7%	10 846	1 001
2003	93,7%	17 312	698
2004	93,7%	18 293	2 363
2005	93,7%	14 660	5 424
2006	93,7%	17 623	13 865
2007	93,7%	28 513	27 391
		R	50 389

D'après tableau ci-dessus, la réserve selon méthode Loss Ratio simple est de 50 389. Cela dit, on remarque que les résultats obtenus au niveau des années 1998 et 2000 sont incohérents ; cela peut être dû à la force de l'hypothèse de base de la méthode.

III.2 La méthode du Loss Ratio complémentaire

Cette méthode consiste à supposer que les Loss Ratios L_i ne varient que sensiblement. Donc, avant d'entamer le calcul, on doit d'abord estimer une valeur commune pour les L_i . Pour cela, on détermine un Loss Ratio $L(j)$ pour chaque délai de règlement j grâce à la formule suivante :

$$L(j) = \frac{\sum_{i=0}^{n-j} C_{ij}}{\sum_{i=0}^{n-j} P_i}$$

L'étape suivante consiste à analyser la convergence de la suite $L(j)$ afin d'estimer le ratio final commun L :

J	0	1	2	3	4	5	6	7	8	9	10
L(j)	5,6%	28,1%	58,5%	76%	84,2%	87,2%	91,3%	98,7%	100,8%	116,6%	93,7%

On remarque que la suite $L(j)$ ne converge pas. Cela remet en cause l'hypothèse de base sur les L_i .

Cependant, il existe une alternative pour le calcul du ratio final. Il s'agit de faire une régression linéaire simple de $L(j)$ en fonction de j et on retient la valeur de $L(j)$ pour $j=10$ en guise de Loss ratio.

Après avoir fait la régression linéaire simple, on trouve l'équation:

$L(j) = 0,088 j + 0,321$ et donc la valeur prévue du loss ratio en l'année 10 est $L=L(10)=120,10\%$

Ensuite, on calcule la provision par les formules ci-dessous :

- $S_i = L * P_i$
- $R_i = S_i - C_{i,n-i}$
- $R = \sum_{i=0}^n R_i$

Ainsi on obtient le tableau suivant:

	P_i	L	$C_{i,n-i}$	S_i	R_i
1997	12 860	102,10%	12 045	13 130	1 085
1998	10 026	102,10%	14 643	10 237	-4 406
1999	14 138	102,10%	10 669	14 435	3 766
2000	12 086	102,10%	11 902	12 340	438
2001	12 856	102,10%	9 132	13 126	3 994
2002	11 580	102,10%	9 845	11 823	1 978
2003	18 483	102,10%	16 614	18 871	2 257
2004	19 531	102,10%	15 930	19 941	4 011
2005	15 652	102,10%	9 236	15 981	6 745
2006	18 815	102,10%	3 758	19 210	15 452
2007	30 442	102,10%	1 122	31 081	29 959
					65 279

La provision pour sinistres trouvée par le biais de cette méthode est 65 279 MDH.

Pour vérifier l'hypothèse de base, à savoir la stabilité relative des Loss Ratios, on utilise le tableau suivant:

	C_{in}	P_i	L_i
1997	12 045	12 860	93,66%
1998	14 648	10 026	146,10%
1999	10 682	14 138	75,56%
2000	12 158	12 086	100,60%
2001	9 539	12 856	74,20%
2002	10 714	11 580	92,52%
2003	19 047	18 483	103,05%
2004	20 564	19 531	105,29%
2005	15 513	15 652	99,11%
2006	12 606	18 815	67,00%
2007	17 739	30 442	58,27%

Effectivement, On constate que les L_i varient considérablement d'année en année; cela explique bien la non-convergence de la série $L(j)$ ainsi que les valeurs aberrantes trouvées au niveau de quelques réserves annuelles.

IV. Comparaison des résultats issus de chacune des méthodes déterministes et d'inventaire

Avant tout développement, le mieux serait de dresser un tableau récapitulatif de l'ensemble des réserves estimées par les méthodes et les techniques vues précédemment.

	Rglmt	C.L Standard	C.L. Moyenne	C.L.3 derniers	C.L. dernier	Lnd Chain	Lnd Pivot	L.R simple	L.R cml
1997	3919	0	0	0	0	-	-	0	1085
1998	1252	4,8	4,8	4,86	4,86	4,9	4	-5252	-4406
1999	952	13,1	13,2	13,24	12,3	15,6	16	2573	3766
2000	963	256	237,3	237,38	21,54	224	263	-582	438
2001	1124	406,7	379,3	404,49	33,47	40667,6	530,25	2909	3994
2002	2955	869,1	841,5	824,31	583,07	338,6	984,85	1001	1978
2003	9786	2433,2	2390,6	2368,52	1704,03	4194,7	1527,18	698	2257
2004	6878	4633,9	4684,9	4391,03	3052,97	5829,7	4939,61	2363	4011
2005	9573	6276,9	6294,5	6074,93	5964,86	6777,6	4939,61	5424	6745
2006	12580	8847,7	8859,5	8746,18	7650,99	3443,3	9065,17	13865	15452
2007	22255	16616,8	16729,6	16113,09	11980,24	2110,5	12294,97	27391	29959
	72237	40358,6	40435,6	39178,04	31008,32	63606,5	34564,64	50389	65279

La nature préventive des méthodes réglementaires traduit les mesures prises pour protéger les droits de l'assuré. Cette optique protectrice fait que la provision préconisée par les autorités de contrôle soit la plus élevée avec un montant de 72 237 MDH

La méthode de Chain Ladder standard et ses variantes donnent, quant à elles, des estimations bien plus optimistes, de l'ordre de 37 000 MDH.

La technique de London Pivot donne un résultat très optimiste. Pour ce qui est du reste des méthodes, elles donnent une estimation plus élevée que celle de Chain Ladder mais moins importante que la réserve réglementaire. En particulier, Le modèle du Loss Ratio complémentaire est le plus pessimiste dans la mesure où il prévoit une réserve dépassant celle estimée par Chain Ladder Standard de 24 920,4 MDH.

Pour plus de précision quant à la dispersion de l'ensemble des réserves calculées jusqu'ici, on présente les caractéristiques suivantes :

Moyenne	48 561,9
Ecart-type	14 979,52
Coefficient de variation	30,85%

On dispose ainsi d'un jeu de provisions estimées. Le choix de la provision ultime dont on tiendra compte dépend de plusieurs considérations comme la définition de la problématique traitée. En effet, selon l'objectif ciblé, la réserve choisie peut-être la moyenne de toutes les estimations, la moyenne tronquée de la plus petite ou de la plus élevée... ou encore reposer sur l'expérience de l'actuaire.

Conclusion

Le large recours et l'universalité des méthodes de Chain Ladder se justifient uniquement par leur simplicité de mise en oeuvre. En effet, Chain Ladder doit sa popularité dans le monde de l'actuariat à sa capacité de manipulation des données triangulaires rien qu'en faisant appel à une régression linéaire simple.

Par ailleurs, bon nombre de méthodes déterministes font de la régression linéaire un pilier majeur. Ainsi, ces méthodes supposent que les éléments du triangle de liquidation revêtent une certaine régularité, chose qui est nuancée dans le monde réel. En effet, plus la branche étudiée est à déroulement lent, plus il devient très difficile d'affirmer de telles hypothèses et de baser ses calculs sur des régressions simples. Cela peut engendrer de grands risques de paramétrisation entraînant des estimations moins robustes.

De ce fait, quoi que simples, les méthodes déterministes présentent plusieurs inconvénients qui limitent leur performance. En voici quelques uns :

- ❑ Elles ne sont habilitées à détecter les irrégularités potentielles du triangle étudié.
- ❑ Elles rendent impossible toute prise en compte d'un changement de jurisprudence, une inflation non-constante, une innovation quant à la gestion de la sinistralité...
- ❑ Elles ne permettent pas la modélisation par une loi de probabilité pour la provision globale.

Ces problèmes et bien d'autres poussent l'actuaire à chercher une méthodologie plus rigoureuse. C'est pourquoi et afin de bien cerner les incertitudes inhérentes à la procédure de provisionnement, il est nécessaire d'avoir recours aux méthodes stochastiques. En effet, ces dernières permettent, en plus d'une modélisation probabiliste de la réserve ultime, de quantifier la marge d'erreur existante pour chaque estimation faite.

Partie III. Les méthodes stochastiques pour le calcul des provisions

Cette section est dédiée aux méthodes stochastiques pour le calcul de la PSAP. Elle est subdivisée en plusieurs chapitres, chacun traitant d'un modèle stochastique à part.

Chaque chapitre est fidèle à une certaine logique dans la construction du modèle. En effet, les hypothèses de base viennent en premier lieu, la construction du modèle et l'estimation de ces paramètres vient ensuite permettre le calcul de la provision ultime et des risques d'estimation qui s'y rapportent. La phase finale est sans doute la validation du modèle par la vérification du degré de la véracité des hypothèses énoncées au tout début de la modélisation.

Le premier chapitre de cette partie se rattache au modèle de Mack. Le second explicite l'application des modèles linéaires généralisés à la problématique du provisionnement.

Introduction

Le recours aux techniques stochastiques pour le calcul des provisions en assurance est très recommandé. Ces méthodes requièrent une application de notions probabilistes et statistiques avancées. De ce fait, elles relèvent des compétences de l'actuaire.

L'adoption des modèles stochastiques en tant qu'outil efficace d'estimation de la réserve pour sinistre vient en réponse aux normes modernes d'évaluation du patrimoine d'une compagnie d'assurances. C'est dans ce sens qu'on parle du principe du « best estimate » des engagements techniques...

Il existe une large variante des méthodes stochastiques pour le calcul des provisions. Chacune de ces méthodes diffère par ses principes de base, ses hypothèses et ses étapes. On peut ainsi distinguer : les méthodes récursives, les méthodes factorielles...

Les techniques probabilistes viennent certainement pallier à quelques lacunes présentes dans l'application des méthodes déterministes. En effet, l'usage des méthodes de calcul à caractère stochastique permet entre autres de quantifier quelques erreurs de prédiction, d'estimer plusieurs paramètres de provisionnement appropriés au type d'interprétation que l'on veut établir... Le plus important est qu'une méthode stochastique donne généralement une explication logique des procédés des techniques déterministes.

Classiquement, une technique stochastique part d'une ou plusieurs hypothèses faites sur la nature des données pour aboutir à un ensemble d'estimations des paramètres de provisionnement convoités. Les données, à savoir les éléments du triangle de liquidation sont pris pour des variables aléatoires. On passe généralement par l'attribution d'une distribution à la variable aléatoire provision notée R . Il se pourrait également que le modèle soit non paramétrique dans la mesure où il n'a pas recours à la loi de R . Bien entendu, une méthode stochastique s'achève sur une vérification des hypothèses de départ pour contrôler la robustesse du modèle, notamment grâce à une analyse des résidus...

Dans ce qui suit, nous allons essayer d'appliquer les méthodes stochastiques qui nous semblent les plus répandues et les plus pertinentes compte tenu des données dont on dispose.

Chapitre I. Le modèle de Mack

De manière très simple, le modèle de Mack n'est autre que le formalisme probabiliste qui justifie la technique déterministe Chain Ladder Standard. L'élaboration de ce modèle est plus ou moins récente ; elle remonte à l'année 1993. Elle est tributaire du grand souci des chercheurs de fournir un modèle stochastique dont les résultats coïncident avec ceux de la méthode de Chain Ladder.

Comme nous allons voir par la suite, ce modèle revêt un aspect conditionnel dans le sens où l'on construit les éléments manquants tout en tenant compte du triangle supérieur. De surcroît, ce modèle ne fait aucune hypothèse quant à la distribution qui s'ajuste aux données et par conséquent sur celle qui sied à la variable aléatoire provision globale R.

I. Les hypothèses sous-jacentes au modèle

Le modèle de Mack repose sur quelques hypothèses de départ. Elles sont au nombre de trois et peuvent être présentées ainsi :

- H1 : Hypothèse sur l'indépendance entre les lignes du triangle de liquidation. Ainsi, C_{ij} est indépendant de $C_{i'j}$ pour $i \neq i'$. Cette hypothèse peut être remise en question, notamment en cas d'un grand changement dans la gestion des sinistres d'une année à une autre...
- H2 : Le processus $(C_{ij})_{j>0}$ est markovien. On introduit les facteurs de développement déjà vus dans la méthode de Chain Ladder Standard f_j et on a, pour chaque $j=0, \dots, n-1$ un f_j tel que pour tout i allant de 0 à n :

$$(*) E(C_{i,j+1}/H_{ij}) = f_j * C_{ij}$$
 H_{ij} étant le vecteur des éléments du triangle des paiements pour une année d'origine i à gauche de $C_{i,j+1}$ ou encore $H_{ij} = (C_{i0}, \dots, C_{ij})$. Si l'on veut reformuler l'hypothèse ne tenant compte que des éléments du d. triangle, on sera amené à multiplier l'équation (*) par $\frac{1}{C_{ij}}$, chose qui est correcte puisque C_{ij} est connu et donc n'est pas une variable aléatoire. Ainsi, nous aurons directement : $E(f_{ij}/H_{ij}) = f_j$.
- H3 : Hypothèse relative à la volatilité des données. Le modèle suppose que pour une année d'origine définie nous avons, pour tout $j=0, \dots, n-1$ on a un $(\sigma_j)^2$ tel que $V(C_{i,j+1}/H_{ij}) = (\sigma_j)^2 * C_{ij}$ ou encore, si on multiplie cette même équation des deux côtés par $(\frac{1}{C_{ij}})^2$, on obtient $V(f_{ij}/H_{ij}) = (\sigma_j)^2 / C_{ij}$.

Remarque : Ces trois hypothèses conduisent au modèle hypothétique suivant :

$C_{i,j+1} = f_j C_{ij} + \sigma_j \sqrt{C_{ij}} + \varepsilon_{ij}$ avec ε_{ij} i.i.d et centrés. Ce modèle donne l'occasion pour estimer ses coefficients à l'aide des moindres carrés pondérés. Dans ce cas, on sera amené à minimiser la somme des carrés des écarts pondérés $\sum_{i=0}^{n-j-1} \frac{1}{C_{ij}} (C_{i,j+1} - f_j C_{ij})^2$.

II. L'estimation des paramètres du modèle

Après avoir exposé les hypothèses du modèle, il convient d'explicitier par la suite les résultats qui en découlent.

Notons $T = [C_{ij} / i+j \leq n]$ le triangle supérieur ou bien le triangle des valeurs observées. Les éléments diagonaux sont également compris dans T.

On a donc : $E(C_{in}/T) = f_{n-1} * f_{n-2} * \dots * f_{n-i} * C_{i,n-i}$. De façon générale, pour $i+j > n$, $E(C_{ij}/T) = f_{j-1} * f_{j-2} * \dots * f_{n-i} * C_{i,n-i}$.

Egalement, nous avons pour tout $j=0 \dots n-1$, un estimateur sans biais de f_j est \hat{f}_j tel que : $\hat{f}_j = \frac{\sum_{i=0}^{n-j-1} C_{i,j+1}}{\sum_{i=0}^{n-j-1} C_{ij}}$. Les \hat{f}_j sont non corrélés.

L'absence du biais permet de conclure qu'à partir de l'expression des espérances déjà établie : $\widehat{C}_{ij} = \widehat{f}_{j-1} * \widehat{f}_{j-2} * \dots * \widehat{f}_{n-i} * C_{i,n-i}$ pour $i+j > n$ ou encore $\widehat{C}_{in} = \widehat{f}_{n-1} * \widehat{f}_{n-2} * \dots * \widehat{f}_{n-i} * C_{i,n-i}$.

Par extension de ce constat, on écrit encore $\widehat{R}_i = \widehat{C}_{in} - C_{i,n-i}$ et donc $\widehat{R} = \sum_{i=0}^n \widehat{R}_i$ au lieu des expressions avec les espérances conditionnelles sur T.

L'application numérique des résultats ci-dessus donne les mêmes aboutissements chiffrés que ceux de la méthode de Chain Ladder standard.

Un estimateur non-biaisé de $(\widehat{6}_j)^2$ est : $(\widehat{6}_j)^2 = \frac{1}{n-j-1} \sum_{i=0}^{n-j-1} C_{ij} \left(\frac{C_{i,j+1}}{C_{ij}} - \widehat{f}_j \right)^2$

Pour le cas à problème de $j=n-1$, où l'estimateur proposé ne donne plus de valeur valable, Mack a proposé la formule alternative suivante :

$$(\widehat{6}_{n-1})^2 = \min \left(\left(\frac{\widehat{6}_{n-2}^2}{\widehat{6}_{n-3}^2} \right)^2, \widehat{6}_{n-2}^2, \widehat{6}_{n-3}^2 \right)$$

j	0	1	2	3	4	5	6	7	8	9
$(\widehat{6}_j)^2$	1316.34	131.78	39.53	20.75	2.05	2.82	11.05	7.06	4.51	0

III. L'estimation des erreurs de prédiction

III.1 Le formalisme mathématique :

Selon le besoin, on sera amené à calculer un type précis d'erreur de prédiction ou d'estimation. Les plus courantes de ces erreurs sont la moyenne quadratique des écarts de prédiction MSEP et l'erreur standard de prédiction SEP ou encore l'erreur standard relative.

De façon générale, si $\widehat{\pi}$ est l'estimation du paramètre π , on définit les erreurs relatives à cette estimation à travers les formules qui suivent:

- $MSEP(\widehat{\pi}) = E(\widehat{\pi} - \pi)^2$
- $SEP(\widehat{\pi}) = \sqrt{MSEP(\widehat{\pi})}$
- $SER(\widehat{\pi}) = \frac{SEP(\widehat{\pi})}{\widehat{\pi}}$

Le plus important est, bien entendu, de pouvoir calculer les erreurs d'estimation de la provision globale. Pour ce faire, on doit calculer l'erreur d'estimation de la provision pour chaque année d'origine. Ainsi, si on pose $\widehat{C}_{i,n-i} = C_{i,n-i}$ on aura les deux résultats suivant :

1. $\widehat{MSEP}(\widehat{R}_i) = \widehat{C}_{in}^2 \sum_{j=n-i}^{n-1} \frac{(\widehat{\sigma}_j)^2}{(\widehat{f}_j)^2} \left[\frac{1}{\widehat{C}_{ij}} + \frac{1}{\sum_{k=0}^{n-j-1} C_{kj}} \right]$
2. $\widehat{MSEP}(\widehat{R}) = \sum_{i=1}^n \left[\widehat{MSEP}(\widehat{R}_i) + \widehat{C}_{in} \left(\sum_{k=i+1}^n \widehat{C}_{kn} \right) \sum_{j=n-i}^{n-1} \frac{2(\widehat{\sigma}_j)^2}{(\widehat{f}_j)^2 \sum_{k=0}^{n-j} C_{kj}} \right]$

Remarque : $\widehat{R}_i = \widehat{C}_{in} - C_{i,n-i}$ et $R_i = C_{in} - C_{i,n-i}$, donc $\widehat{R}_i - R_i = \widehat{C}_{in} - C_{in}$. Les biais des deux estimateurs étant ainsi égaux, on conclut directement que $\widehat{MSE}(\widehat{R}_i) = \widehat{MSE}(\widehat{C}_{in})$.

III.2 Application sur les estimations des réserves annuelles

	$\widehat{MSEP}(\widehat{R}_i)$	$\widehat{Se}(\widehat{R}_i)$	$\widehat{SeR}(\widehat{R}_i)$
1997	-	-	-
1998	0,55495177	0,74495085	0,15314405

1999	3,45088363	1,85765541	0,141246
2000	55123,9268	234,784852	0,91707546
2001	124355,686	352,641015	0,8670364
2002	166696,647	408,285007	0,46977868
2003	380354,468	616,728845	0,25346405
2004	828925,986	910,453725	0,19647324
2005	1184588,69	1088,38812	0,17339376
2006	2274917,17	1508,28285	0,17047135
2007	19889861,7	4459,80511	0,26839028

La lecture directe du tableau des erreurs ci-dessus permet d'affirmer que le niveau du risque d'estimation des réserves est largement acceptable. Cela est surtout constatable à partir des valeurs prises par l'erreur standard relative des diverses réserves annuelles qui ne dépassent pas les 27% que pour les trois années : 2000, 2001 et 2002.

On peut effectuer un choix quant à la distribution de la variable aléatoire réelle R_i . Cela dit, ce choix demeure plus ou moins arbitraire : Il se pourrait bien que la loi votée ne soit pas aussi ajustée que cela aux données. On préfère donc travailler à l'aide d'une approche non paramétrique.

Cependant, pour pouvoir estimer un intervalle de confiance pour les différentes réserves, on suppose une distribution normale et on a :

$$IC_{5\%}(\widehat{R}_i) = [\widehat{R}_i - 2 * SEP(\widehat{R}_i), \widehat{R}_i + 2 * SEP(\widehat{R}_i)]$$

Le tableau suivant fournit les différents intervalles de confiance au seuil de 5% pour toutes les réserves annuelles:

i	R_i	borne inf	borne sup
1997	-	-	-
1998	4,86438003	3,40427637	6,324483697
1999	13,1519156	9,51091102	16,79292023
2000	256,014758	-204,163552	716,1930676
2001	406,719967	-284,456423	1097,896358
2002	869,100752	68,8621372	1669,339366
2003	2433,2005	1224,41196	3641,989031
2004	4633,98347	2849,49417	6418,472774
2005	6276,97399	4143,73328	8410,214703
2006	8847,72041	5891,48602	11803,9548
2007	16616,8652	7875,64714	25358,08319

Remarque : L'hypothèse faite sur la distribution des R_i nous fait tomber dans le cas où une partie de l'intervalle construit est négative puisque le domaine de la loi normale est l'ensemble des nombres réels. Cependant, on pourrait détourner ce problème en munissant R_i d'une loi Log-normale.

III.3 Application sur l'estimation de la réserve globale

A l'instar de l'étude faite pour les réserves annuelles, on analyse les erreurs pour l'estimation de la réserve globale R et on construit l'intervalle de confiance qui s'y rapporte :

\hat{R}	MSEP(\hat{R})	Se(\hat{R})	SeR(\hat{R})	borne inf	borne sup
40358,5953	29176853,23	5401,56026	0,133839154	29555,47479	51161,7158

L'estimation de la réserve globale R est très satisfaisante. L'erreur standard relative étant de 13,38%, on peut largement affirmer que l'on est en présence d'une bonne approximation de la valeur réelle de la provision pour sinistres à constituer.

IV. La validation des hypothèses du modèle

La deuxième hypothèse du modèle est vérifiée de la même manière que pour la méthode de Chain Ladder Standard. Autrement dit, L'hypothèse H2 peut être vue comme une hypothèse de régression linéaire entre les deux séries successives C_j et C_{j+1} . Cette droite ne doit pas avoir de constante. Sa courbe doit donc passer par l'origine. Ainsi, on peut se contenter d'une vérification graphique comme ce fût le cas lors de l'application de la méthode Chain Ladder (Annexe 1).

IV.1 La vérification de l'hypothèse d'indépendance H1 par le test non paramétrique de la médiane

Pour tester l'hypothèse H1, on procède par un test non paramétrique se basant sur le D.triangle et les médianes des éléments des diagonales. En effet, ce test consiste à calculer la médiane pour chaque colonne du D.triangle.

Notons par A_j le vecteur qui regroupe les éléments de la $j^{\text{ème}}$ diagonale du D.triangle pour $j=0 \dots n$.

Il faut ensuite compter le nombre des éléments du vecteur A_j supérieurs à la médiane déjà calculée ainsi que le nombre des éléments du vecteur A_j qui lui sont inférieurs. On note respectivement ces deux nombres par G_j et P_j .

On note également $Z_j = \min(G_j, P_j)$ et on calcule l'espérance et la variance de Z_j par les formules suivantes :

- $E(Z_j) = \frac{n_j}{2} - \binom{n_j-1}{m_j} \frac{n_j}{2^{n_j}}$
- $\text{Var}(Z_j) = \frac{n_j(n_j-1)}{4} - \binom{n_j-1}{m_j} \frac{n_j((n_j-1))}{2^{n_j}} + E(Z_j) - E(Z_j)^2$

Avec $n_j = G_j + P_j$ et $m_j = \frac{n_j-1}{2}$

Prenons Z la somme des Z_j pour $j=1 \dots n$. Donc l'espérance de Z est la somme des espérances des Z_j et la variance de Z est, sous l'hypothèse d'indépendance, égale est la somme des variances des Z_j .

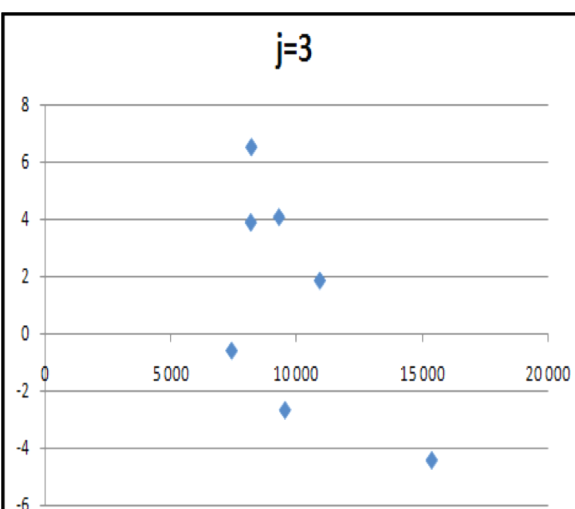
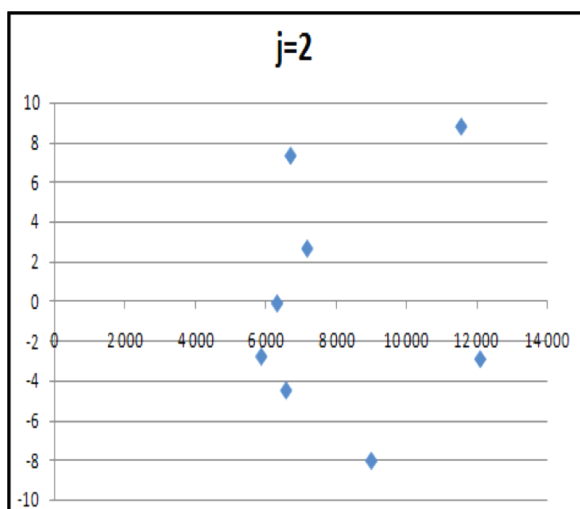
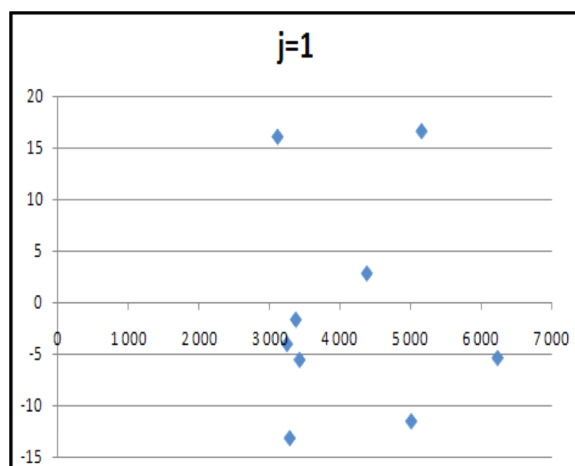
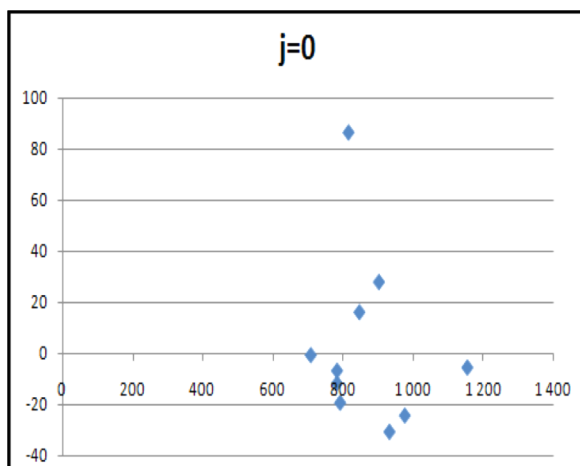
On suppose que Z suit une distribution normale et on accepte l'hypothèse au seuil 5% si Z appartient à l'intervalle $IC_{5\%}(Z) = [E(Z) - 2\sqrt{\text{VAR}(Z)}, E(Z) + 2\sqrt{\text{VAR}(Z)}]$.

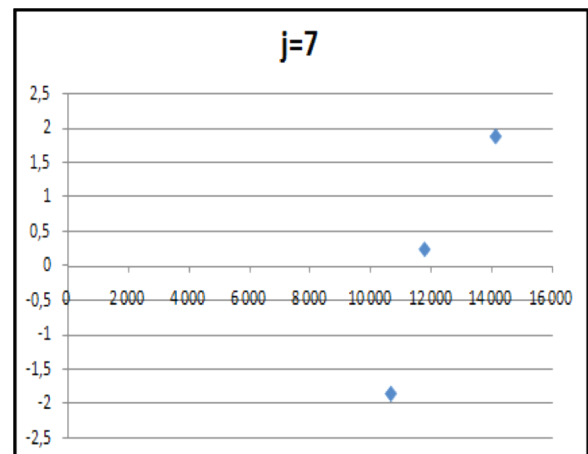
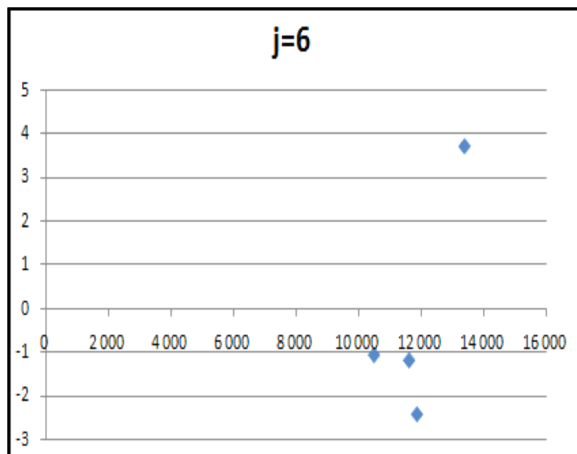
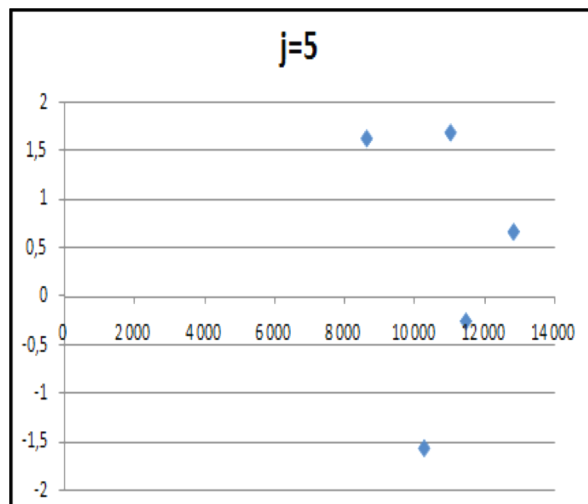
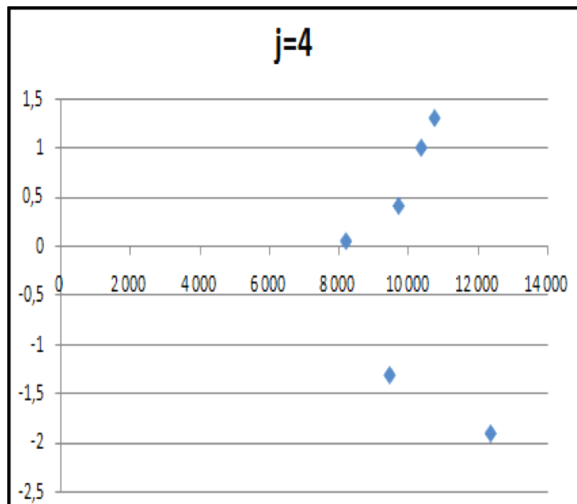
DIAG	G_j	P_j	Z_j	$E(Z_j)$	$\text{var}(Z_j)$
1	2	0	0	0,5	0,25
2	3	0	0	0,75	0,1875
3	0	4	0	1,25	0,4375
4	3	1	1	1,25	0,4375
5	3	3	3	2,0625	0,62109375
6	3	3	3	2,0625	0,62109375
7	3	3	3	2,0625	0,62109375
8	6	2	2	2,90625	0,80371094
9	2	7	2	3,26953125	0,73594666
total			14	16,11328125	4,71543884

L'intervalle de confiance est donc [11,770269 ; 20,4562935] et $Z=14$ appartient à ce dernier. On accepte donc l'hypothèse de la non-signification de l'effet de d'année calendaire.

IV.2 La vérification de l'hypothèse de volatilité H3 grâce aux résidus normalisés

A ce stade, on procède à la vérification de l'hypothèse H3. Pour ce faire, on a recours aux résidus normalisés $r_{ij} = \frac{c_{i,j+1} - \hat{f}_j c_{ij}}{\sqrt{c_{ij}}}$ pour $j=0 \dots n-1$ et $i=0 \dots n-j-1$. On trace par la suite le graphe de ces résidus. Pour valider l'hypothèse, la courbe obtenue ne doit pas laisser entrevoir l'existence d'une composante déterministe et surtout une tendance.





Remarque : On arrête la génération des graphes pour $j=7$ car à partir de ce stade où le diagramme regroupe trois points, on ne peut plus juger de la tendance du nuage créé.

Dans chacun des nuages de points représentés ci-dessus, on remarque un parfait éparpillement et donc l'absence de toute tendance, excepté le dernier graphe où le pseudo-alignement des points pourrait s'assimiler à une tendance linéaire. Conséquemment, on accepte l'hypothèse H3.

Chapitre II. Le calcul de provisions par la modélisation GLM

I. Présentation générale des modèles GLM

Les modèles linéaires généralisés ont été définis pour la première fois par John Nelder et Robert Wedderburn en 1972. Ils sont venus pallier à l'étroitesse des modèles linéaires.

En effet, la classe des GLM est une extension des modèles linéaires traditionnels qui permet à la moyenne d'une population de dépendre d'un estimateur linéaire via une fonction non linéaire dite *fonction lien*. Ainsi, le choix d'une fonction lien adéquate et d'une distribution pour la variable réponse sont inévitables à la construction d'un modèle linéaire généralisé. De surcroît, la modélisation GLM est paramétrique dans le sens où la variable expliquée est supposée munie d'une loi de probabilité. Plus exactement, elle est prétendue appartenir à la famille des lois exponentielle. Les modèles GLM autorisent également la variance de chaque mesure à être une fonction de sa valeur prévue.

Les avantages précités et d'autres encore font des modèles linéaires généralisés l'un des choix de modélisation statistique les plus populaires, notamment en matière de provisionnement.

II. Formalisme mathématique et construction du modèle

II.1 L'hypothèse du modèle

La modélisation en GLM se base sur l'hypothèse que l'échantillon est constitué de variables aléatoires indépendantes. Cette hypothèse est vérifiable grâce à une multitude de tests statistiques.

II.2 Les composantes du modèle

La construction d'un modèle linéaire généralisé requiert la disposition puis l'assemblage de deux composantes:

- La partie linéaire du modèle qui se résume en l'estimateur linéaire: $\omega = X\hat{\beta}$ (Voir annexe2).
- La fonction lien notée g qui décrit l'obtention de \hat{Y} , la valeur prévue de Y , à partir de l'estimateur linéaire : $g(\hat{Y}) = \omega$.

II.3 La distribution de la variable réponse et la famille de lois exponentielle

En modélisation GLM, la distribution de la variable réponse Y doit appartenir à la famille de lois exponentielle. La forme générale de la fonction de masse pour cette famille de lois s'écrit :

$$f(y, \theta, \Phi) = \text{Expo} \left[\frac{(\theta y - b(\theta))}{\Phi} + c(y, \Phi) \right]$$

y étant une réalisation de la variable Y .

θ : est un paramètre réel dit paramètre naturel.

Φ : il s'agit du paramètre de dispersion. Il est connu et doit être positif.

Pour ce qui est des fonctions b et c , elles sont spécifiques à chaque loi. Eventuellement, la fonction b doit être régulière pour permettre le calcul d'un certain nombre de paramètres de la distribution. Il a été démontré en effet que :

- $\mu = E[Y] = b'(\theta)$ et si de plus b^{-1} existe, alors on a $\theta = b'^{-1}(\mu)$.
- $V(Y) = \Phi b''(\theta) = \Phi V(\mu)$.

De façon générale, il est possible de déduire l'expression de tous les moments de Y rien qu'en dérivant la fonction génératrice des moments :

$$M_Y(s) = \text{Exp}\left(\frac{1}{\Phi} [b(\theta + sy) - b(\theta)]\right)$$

II.3.1 La loi de poisson $P(\lambda)$

La loi de poisson décrit le dénombrement d'évènements dans un intervalle. Si λ est le nombre d'occurrences moyen ($\lambda > 0$) et Y la variable discrète à modéliser, on aura :

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

Par passage à la forme générale de la famille exponentielle, on aura:

$$P(Y=y) = e^{\left(\frac{y \ln \lambda - \lambda}{\Phi} - \ln(y!)\right)}$$

Par analogie, on identifie $\theta = \ln \lambda$, $b(\theta) = e^\theta$, $\Phi = 1$ et $c(y, \Phi) = -\ln(y!)$ et on en déduit l'espérance et la variance de la variable étudiée :

$$\begin{cases} \mu = E[Y] = b'(\theta) = \lambda \\ V(Y) = \Phi b''(\theta) = \Phi V(\mu) = \lambda \end{cases}$$

II.3.2 La loi Gamma $G\left(v, \frac{v}{\mu}\right)$

La loi Gamma permet de modéliser une grande variété de phénomènes pour des grandeurs positives. Si Y suit une loi Gamma de paramètres a et b, alors la fonction de masse s'écrit comme suit :

$$f(y, a, b) = \frac{b^a}{\Gamma(a)} y^{a-1} e^{-by} \text{ pour } x > 0$$

Une simplification dans le cadre de l'écriture sous la forme exponentielle consiste à poser $a=v$ et $b=\frac{v}{\mu}$ et on réécrit la fonction de densité:

$$f(y) = e^{\left[\left(-\frac{v}{\mu} - \ln \mu\right)v + c(y, v)\right]}$$

$$\text{Par identification, on trouve } \begin{cases} \Phi = \frac{1}{v} \\ \theta = -\frac{1}{\mu} \\ E[Y] = \mu \\ V(\mu) = \mu^2 \end{cases}$$

III. Application des modèles GLM aux triangles de liquidation

Dans le cadre des méthodes stochastiques dédiées au calcul des provisions pour sinistres, la modélisation GLM tente de fournir une estimation ponctuelle et par intervalle de confiance pour les différentes réserves annuelles ou globale. Ainsi, elle sélectionne le modèle probabiliste le plus adéquat aux incréments de paiement du triangle supérieur. Elle permet également le calcul des différents risques d'estimation ainsi que les mesures de la qualité d'ajustement du modèle choisi.

Dans cette section, on appliquera l'approche GLM sur les éléments du triangle de liquidation supposés aléatoires.

III.1 Les différents éléments du modèle

La modélisation stochastique GLM part de la supposition que les incréments de paiements du rectangle de liquidation sont des variables aléatoires. On les note $(Y_{ij})_{i,j=0,\dots,n}$.

$i \backslash j$	0	1				i		l			n
0	y_{11}	y_{12}									
1	y_{21}	y_{22}									
i						y_{ii}					
k								Y_{kl}			
n											Y_{nn}

Les éléments du triangle supérieur, c'est-à-dire de la zone rose ont été en effet observés. Il s'agit donc de réalisations bien connues qu'on note $(y_{ij})_{i+j \leq n}$.

Remarque : Le triangle supérieur contient également les règlements calendaires (éléments de la diagonale): quelque soit l'année d'origine, le paiement de la sinistralité a été effectué en l'année n. Ainsi, la somme des éléments diagonaux correspond au montant payé par l'assureur au cours de l'année comptable n toutes années de survenance confondues.

La zone bleue, quant à elle, regroupe les règlements restant à effectuer pour les années de survenance de 0 à n. Les éléments du triangle inférieur sont donc des variables aléatoires et on les note $(Y_{ij})_{i+j > n}$.

On suppose par la suite que les $(Y_{ij})_{i,j=0..n}$ sont identiquement distribuées. On y rajoute l'hypothèse d'indépendance pour pouvoir utiliser l'approche GLM.

En conséquence, le triangle supérieur, assimilé à un échantillon observé, sert principalement à retrouver les paramètres du modèle. La suite consiste à utiliser ces paramètres pour estimer les règlements futurs suivant le modèle retenu.

Le choix des variables explicatives est la phase suivante de la modélisation. Les variables sélectionnées pour leur influence explicite sur la variable dépendante Y_{ij} sont :

- L'année de survenance i: elle sera paramétrée par un coefficient α_i . Comme nous avons (n+1) années de survenance, il est normal d'avoir (n+1) paramètres chacun renvoyant à une année d'origine bien précise.
- Le délai de règlement j: sera représenté par un autre paramètre β_j . De même que pour l'année de survenance, le délai de règlement donne lieu à (n+1) autres paramètres.
- L'année calendaire (i+j): pour des données déflatées, chose que nous supposons ici, l'effet "année calendaire" est supposé constant. C'est ainsi que le paramètre correspondant à cette variable sera tenu pour constant μ .

La construction du modèle nécessite également la donnée d'une fonction f qui lie la moyenne $\mu_{ij} = E[Y_{ij}]$ à l'ensemble des paramètres de régression prédéfinis :

$$\mu_{ij} = g(\alpha_i, \beta_j, \mu)$$

Par référence à la méthode Chain Ladder et pour l'interprétation du paramétrage qu'il permet, le lien exponentiel est classiquement pris et on obtient:

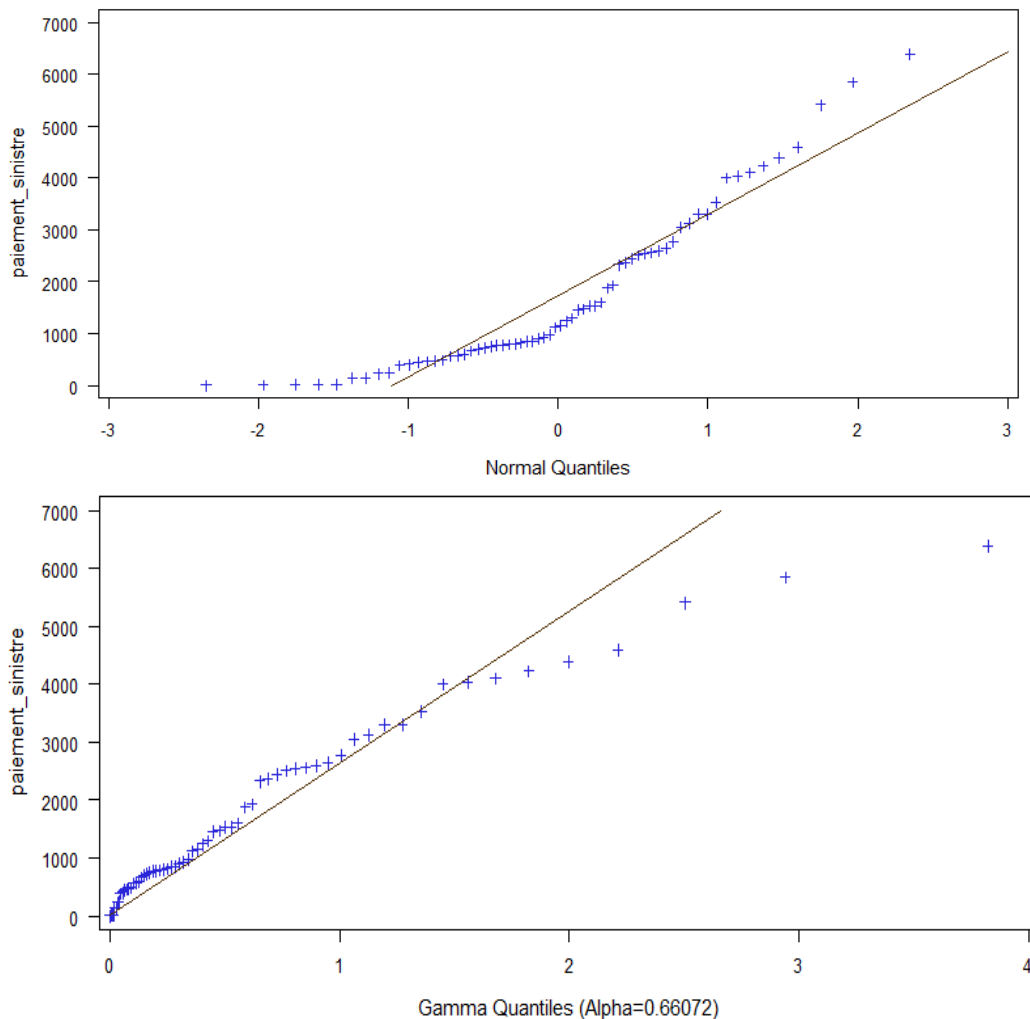
$$\mu_{ij} = e^{\mu + \alpha_i + \beta_j}$$

III.2 Le choix de la distribution de la variable paiement des sinistres

La première étape consiste à choisir la distribution de la variable réponse. Plusieurs tests permettent de s'assurer de la pertinence de tel ou tel choix. Dans notre étude, nous aurons recours surtout aux diagrammes Quantile-Quantile et à la comparaison des distances de Kolmogorov-Smirnov.

Les distributions sur lesquelles nous focaliserons notre analyse sont la loi normale et la loi gamma. Ces deux lois feront l'objet d'une petite analyse de la qualité d'ajustement avant d'être choisie. Nous avons également décidé de travailler avec la loi de poisson pour pouvoir vérifier le constat selon lequel ses résultats convergent vers ceux de Chain ladder standard.

Les graphes suivants sont des Q-Q Plot respectivement pour la loi normale et pour la loi gamma :



Sur le tracé relatif à la loi gamma, les points générés s'alignent assez bien autour de la droite. Ce constat n'est pas évident pour le graphe Q-Q plot de la loi normale. En conséquence, la distribution qui sied le mieux aux données du triangle supérieur est la loi gamma (Voir annexe 3).

La comparaison des distances de Kolmogorov-Smirnov confirme également ce choix : quand $d=0,16567$ pour la loi normale, sa valeur n'est que $0,08949$ en ce qui concerne la loi gamma ; et comme le critère de sélection du modèle est d'avoir la plus petite valeur possible de la distance d , la distribution gamma est choisie (Voir Annexe 4).

Remarque : L'analyse à l'aide de l'outil graphique Q-Q plot ou de la distance de K-S n'est que descriptive. Elle permet d'avoir une première idée sur la meilleure distribution à choisir sans pour autant confirmer définitivement ce choix, d'autant plus que notre jeu de

données est assez réduit (de taille 66). Par la suite, et pour valider le modèle sélectionné, on doit impérativement avoir recours à des tests statistiques plus rigoureux.

III.3 L'estimation et le risque d'estimation

III.3.1 L'estimation des paramètres du modèle et le calcul des réserves

1) Le formalisme théorique

Le premier objectif de cette section est de pouvoir trouver l'estimation du vecteur $((\alpha_i)_{i=0..n}, (\beta_j)_{j=0..n}, \mu)$. Pour cela, la méthode du maximum de vraisemblance est appliquée aux données du triangle supérieur $L((x_{ij})_{i+j \leq n}, (\alpha_i)_{i=0..n}, (\beta_j)_{j=0..n}, \mu)$ dans le but de donner l'estimateur du maximum de vraisemblance $((\hat{\alpha}_i)_{i=0..n}, (\hat{\beta}_j)_{j=0..n}, \hat{\mu})$.

La matrice X du modèle est construite à partir de valeurs binaires (0 et 1) de telle sorte à faire coïncider chaque valeur prévue de la variable à modéliser avec le paramétrage qui lui correspond.

Comme la méthode du maximum de vraisemblance permet l'invariance fonctionnelle, l'estimation des paramètres du modèle conduit directement à l'estimation ponctuelle des différentes valeurs $\hat{\mu}_{ij}$. En effet, la valeur prévue du modèle est donnée par :

$$\hat{\mu}_{ij} = e^{\hat{\alpha}_i + \hat{\beta}_j + \hat{\mu}} \quad \text{pour tout } i=0..n \text{ et } j > n-i$$

Une fois les paiements restant à effectuer estimés, l'étape qui suit consiste à trouver les réserves par année d'origine ainsi que la provision globale à retenir.

Le calcul de la réserve moyenne $\hat{E}(R_i)$ pour l'année de survenance i ne requiert que les données du triangle inférieur. Il s'agit concrètement de sommer tous les éléments de la $i^{\text{ème}}$ ligne du triangle inférieur: $\hat{E}(R_i) = \sum_{j=n-i+1}^n \hat{\mu}_{ij}$.

De même, on déduit la valeur de la provision globale R: $\hat{E}(R) = \sum_{i=0}^n \hat{E}(R_i)$.

2) Application numérique :

Les calculs et analyses qui vont suivre et qui ont été appliqués sur notre base de données ont été élaborés avec le logiciel SAS.

Pour chaque choix de modélisation, c'est-à-dire que ce soit avec la distribution poisson ou la distribution Gamma, l'étape initiale est commune. Elle consiste en la construction de la matrice X. Afin d'illustrer son allure, prenons en les 20 premières lignes :

```

110000000000100000000000
11000000000000100000000000
11000000000000010000000000
11000000000000001000000000
11000000000000000100000000
11000000000000000010000000
11000000000000000001000000
11000000000000000000100000
11000000000000000000010000
11000000000000000000001000
11000000000000000000000100
11000000000000000000000010
11000000000000000000000001
10100000000001000000000000
10100000000000100000000000
10100000000000010000000000
10100000000000001000000000
10100000000000000100000000
10100000000000000010000000
10100000000000000001000000
10100000000000000000100000
10100000000000000000010000
10100000000000000000001000
10100000000000000000000100
101000000000000000000000100

```

a) Le modèle de poisson

Le tableau qui suit regroupe l'ensemble des estimations des paramètres du modèle :

Analyse des valeurs estimées du paramètre de vraisemblance maximum

Paramètre	DDL	Valeur estimée	Erreur type	Intervalle de confiance de Wald à 95 %		Khi-2 de Wald	Pr > Khi-2
Intercept	1	1.7734	0.5011	0.7913	2.7555	12.53	0.0004
A1	1	-0.3871	0.0330	-0.4517	-0.3225	137.97	<.0001
A2	1	-0.1915	0.0327	-0.2556	-0.1273	34.22	<.0001
A3	1	-0.5072	0.0331	-0.5721	-0.4423	234.55	<.0001
A4	1	-0.3778	0.0330	-0.4424	-0.3132	131.35	<.0001
A5	1	-0.6204	0.0333	-0.6857	-0.5551	346.28	<.0001
A6	1	-0.5042	0.0332	-0.5693	-0.4391	230.51	<.0001
A7	1	0.0712	0.0326	0.0073	0.1350	4.77	0.0289
A8	1	0.1478	0.0326	0.0839	0.2116	20.57	<.0001
A9	1	-0.1341	0.0332	-0.1992	-0.0690	16.29	<.0001
A10	1	-0.3416	0.0353	-0.4108	-0.2724	93.51	<.0001
A11	0	0.0000	0.0000	0.0000	0.0000	.	.
B1	1	5.2495	0.5002	4.2691	6.2298	110.14	<.0001
B2	1	6.5614	0.5001	5.5812	7.5416	172.13	<.0001
B3	1	6.7969	0.5001	5.8168	7.7771	184.72	<.0001
B4	1	6.2913	0.5001	5.3111	7.2715	158.24	<.0001
B5	1	5.6832	0.5002	4.7029	6.6636	129.10	<.0001
B6	1	4.9446	0.5004	3.9638	5.9253	97.65	<.0001
B7	1	4.7525	0.5005	3.7715	5.7334	90.16	<.0001
B8	1	4.1742	0.5010	3.1923	5.1561	69.42	<.0001
B9	1	4.0894	0.5014	3.1066	5.0721	66.52	<.0001
B10	1	0.9960	0.5401	-0.0626	2.0546	3.40	0.0652
B11	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale	0	1.0000	0.0000	1.0000	1.0000	.	.

L'erreur standard mesurant le degré de variabilité de l'estimation des paramètres se trouve à un niveau acceptable. On peut par conséquent retenir les valeurs estimées des paramètres du modèle.

Par la suite, on extrait le rectangle des valeurs prévues des incréments de paiements. La dernière colonne du tableau contient la réserve relative aux sinistres survenus en l'année i .

$i \backslash j$	0	1	2	3	4	5	6	7	8	9	10	R_i
0	792	2 327	4 036	2 367	848	666	595	152	246	12	4	0
1	847	3 533	4 597	1 931	1 462	461	561	751	488	12	5	5
2	708	2 545	3 047	1 879	1 539	569	231	144	7	10	4	13
3	933	2 445	3 305	2 599	1 471	720	407	22	241	11	4	256
4	783	2 508	2 561	1 541	809	451	479	206	189	9	3	407
5	783	2 646	3 128	1 602	1 299	387	412	231	212	10	4	869
6	816	5 411	5 847	3 298	1 242	888	733	411	378	17	6	2433
7	1 156	3 998	6 380	4 396	2007	959	791	444	408	18	7	4634
8	903	4 104	4 229	2781	1514	723	597	335	308	14	5	6277
9	977	2 781	3747	2260	1230	588	485	272	250	11	4	8848
10	1122	4166	5273	3180	1731	827	683	383	352	16	6	16617

Ainsi, la valeur prévue de la réserve globale obtenue suite à cette modélisation est: $E(R) = 40\ 359$.

Il est clair que les résultats obtenus par la modélisation poisson sont parfaitement identiques à ceux de la méthode déterministe Chain Ladder Standard avec les mêmes coefficients de développement déjà établis.

Le résultat qu'on vient de vérifier fut établi par Renshaw et Verrall. Il stipule en effet que, si on part d'un triangle de liquidation à incréments positifs, les éléments du triangle

inférieur, les réserves par année d'origine ainsi que la réserve totale sont identiques pour la modélisation de poisson et pour la méthode de Chain Ladder Standard.

b) Le modèle Gamma

Pareil que pour le modèle précédent, on commence par l'estimation des différentes valeurs des paramètres du modèle.

Analyse des valeurs estimées du paramètre de vraisemblance maximum

Paramètre	DDL	Valeur estimée	Erreur type	Intervalle de confiance de Wald à 95 %		Khi-2 de Wald	Pr > Khi-2
Intercept	1	1.6891	0.6041	0.5050	2.8732	7.82	0.0052
A1	1	-0.3028	0.4478	-1.1805	0.5748	0.46	0.4989
A2	1	0.0614	0.4517	-0.8238	0.9466	0.02	0.8919
A3	1	-0.6027	0.4463	-1.4774	0.2719	1.82	0.1768
A4	1	-0.3964	0.4479	-1.2742	0.4814	0.78	0.3761
A5	1	-0.5346	0.4506	-1.4178	0.3486	1.41	0.2355
A6	1	-0.4796	0.4541	-1.3697	0.4104	1.12	0.2909
A7	1	0.0009	0.4621	-0.9048	0.9066	0.00	0.9984
A8	1	0.1633	0.4691	-0.7562	1.0827	0.12	0.7279
A9	1	-0.1221	0.4814	-1.0656	0.8213	0.06	0.7997
A10	1	-0.2481	0.5031	-1.2342	0.7381	0.24	0.6220
A11	0	0.0000	0.0000	0.0000	0.0000	.	.
B1	1	5.3338	0.4478	4.4561	6.2114	141.87	<.0001
B2	1	6.6128	0.4486	5.7335	7.4920	217.30	<.0001
B3	1	6.8442	0.4482	5.9658	7.7225	233.23	<.0001
B4	1	6.3405	0.4504	5.4577	7.2232	198.18	<.0001
B5	1	5.7776	0.4548	4.8862	6.6689	161.39	<.0001
B6	1	5.0038	0.4554	4.1111	5.8964	120.71	<.0001
B7	1	4.7653	0.4589	3.8660	5.6647	107.85	<.0001
B8	1	4.0212	0.4735	3.0932	4.9492	72.13	<.0001
B9	1	3.9028	0.4805	2.9610	4.8445	65.98	<.0001
B10	1	0.9330	0.5025	-0.0519	1.9179	3.45	0.0634
B11	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale	1	6.0808	1.0308	4.3618	8.4772	.	.

Il est constaté que les erreurs standards représentent une part relativement faible des estimations des paramètres. Ces dernières sont par conséquent retenues.

On en vient par la suite au tableau complété des incréments de règlements. Sur la dernière colonne se trouve la valeur prévue des réserves par année de survenance.

I \ j	1	2	3	4	5	6	7	8	9	10	11	R _i
1	792	2 327	4 036	2 367	848	666	595	152	246	12	4	0.0
2	847	3 533	4 597	1 931	1 462	461	561	751	488	12	5.76	5.8
3	708	2 545	3 047	1 879	1 539	569	231	144	7	7.5	2.96	10.5
4	933	2 445	3 305	2 599	1 471	720	407	22	180.4	9.2	3.64	193.4
5	783	2 508	2 561	1 541	809	451	479	176.9	157.1	8.0	3.17	345.3
6	783	2 646	3 128	1 602	1 299	387	393.4	186.9	166.0	8.5	3.35	758.2
7	816	5 411	5 847	3 298	1 242	807.4	636.1	302.2	268.4	13.7	5.42	2033.4
8	1 156	3 998	6 380	4 396	2058.9	949.7	748.2	355.5	315.8	16.2	6.37	4450.7
9	903	4 104	4 229	2717.4	1547.7	713.9	562.4	267.2	237.4	12.1	4.79	6063.2
10	977	2 781	3964.7	2395.8	1364.5	629.4	495.8	235.6	209.3	10.7	4.23	9310.4
11	1 122	4031.5	5081.1	3070.4	1748.8	806.7	635.5	301.9	268.2	13.7	5.41	15963.4

La réserve total au titre de l'année 2007 calculée avec une modélisation Gamma est:
 $E(R) = 39\,134,24$

III.3.2 L'ajustement des modèles et les risques d'estimation

1) Les risques d'estimation

L'incertitude relative à l'estimation des $E[\widehat{R}_i]$ pour toute année de survenance et de $E[\widehat{R}]$ se mesure principalement par la donnée des $V(E[\widehat{R}_i])$ et de $V(E[\widehat{R}])$ et on a pour tout $i=0\dots n$:

$$V(E[\widehat{R}_i]) = \sum_{j=n-i+1}^n V(\widehat{\mu}_{1j}) + \sum_{i=n-j+1}^n \sum_{\substack{k=n-j+1 \\ k \neq i}}^n Cov(\widehat{\mu}_{1j}, \widehat{\mu}_{1k})$$

$$V(E[\widehat{R}]) =$$

Ce qui est contraignant quant à l'usage de cette expression, c'est que le calcul des variances et covariances impliquées dans la formule ci-dessus est souvent hors d'atteinte. Remédier à ce problème requiert le recours à quelques techniques de simulation comme la méthode de bootstrap qui sera bien détaillée dans le chapitre suivant, ou bien l'usage d'une autre alternative asymptotique : la méthode du delta.

En effet, l'application de la méthode du delta conduit à une mesure du risque standard de la provision totale et des provisions annuelles. Egalement, la normalité asymptotique voulue par cette méthode nous conduit à donner les intervalles de confiance pour chacune des réserves précitées :

Le tableau qui suit regroupe l'ensemble des best estimates des provisions par exercice puis de la provision globale au 31/12/2007, pour chacun des deux modèles : Gamma et poisson. Ce tableau est complété par les risques d'estimation correspondants sous la forme de l'erreur standard relative estimée.

i	Modèle de Poisson				Modèle de Gamma			
	$E[\widehat{R}_i]$	SER	Sup	Inf	$E[\widehat{R}_i]$	SER	Sup	Inf
1998	5,8	0,17	3,828	7,772	5	0,11	3,9	6,1
1999	10,5	0,28	4,62	16,38	13	0,23	7,02	18,98
2000	193,4	0,32	69,624	317,176	256	0,08	215,04	296,96
2001	345,3	0,09	283,146	407,454	407	0,12	309,32	504,68
2002	758,2	0,4	151,64	1364,76	869	0,32	312,84	1425,16
2003	2033,4	0,68	-732,024	4798,824	2433	0,64	-681,24	5547,24
2004	4450,7	0,11	3471,546	5429,854	4634	0,54	-370,72	9638,72
2005	6063,2	0,23	3274,128	8852,272	6277	0,23	3389,58	9164,42
2006	9310,4	0,09	7634,528	10986,272	8848	0,12	6724,48	10971,52
2007	15963,4	0,15	11174,38	20752,42	16617	0,1	13293,6	19940,4
Globale	39134,3	0,25	19567,15	58701,45	40359	0,22	22601,04	58116,96

La partie majeure des erreurs standards se trouvent à un niveau acceptable. Les estimations sont par conséquent fiables.

2) La mesure de la qualité d'ajustement

Un modèle statistique est une description mathématique servant à prédire les valeurs prises par une grandeur aléatoire. Il est donc nécessaire, après la construction d'un modèle, de calculer quelques mesures d'ajustement. Ces dernières servent notamment à quantifier l'écart entre les valeurs obtenues suite au modèle et celles appartenant à l'échantillon observé.

Dans le cadre de nos données, nous ciblerons l'étude de l'adéquation de chacun des modèles (poisson et Gamma) aux réalisations de règlements du triangle supérieur. Pour cela, nous aurons recours à quelques concepts statistiques : déviance, résidus...

a) Le formalisme mathématique

Si on note $\widehat{\theta}_{ij} = b^{-1}(\widehat{\mu}_{ij})$ et $\widetilde{\theta}_{ij} = b^{-1}(y_{ij})$, la statistique déviance se définit comme suit : $D = 2 \sum_{i+j \leq n} [y_{ij}(\widetilde{\theta}_{ij} - \widehat{\theta}_{ij}) - (b(\widetilde{\theta}_{ij}) - b(\widehat{\theta}_{ij}))]$.

Vu leur impertinence en GLM, les résidus bruts $r_{ij} = y_{ij} - \widehat{\mu}_{ij}$ sont remplacés par des résidus dits de déviance :

$$r_{ij}^D = \text{sgn}(y_{ij} - \widehat{\mu}_{ij}) \sqrt{2(y_{ij}(\widetilde{\theta}_{ij} - \widehat{\theta}_{ij}) - (b(\widetilde{\theta}_{ij}) - b(\widehat{\theta}_{ij})))}$$

Ainsi, la déviance peut s'écrire en fonction des résidus r_{ij}^D et on a : $D = \sum_{i+j \leq n} (r_{ij}^D)^2$. Pour les distributions considérées, à savoir poisson et gamma, la déviance s'écrit :

Poisson $P(\lambda)$	$D = 2 \sum_{i+j \leq n} \left[y_{ij} \ln \left(\frac{y_{ij}}{\widehat{\mu}_{ij}} \right) - (y_{ij} - \widehat{\mu}_{ij}) \right]$
Gamma $\gamma(v, \frac{v}{\mu})$	$D = 2 \sum_{i+j \leq n} \left[\frac{y_{ij} - \widehat{\mu}_{ij}}{\widehat{\mu}_{ij}} - \ln \left(\frac{y_{ij}}{\widehat{\mu}_{ij}} \right) \right]$

La déviance sert également à construire un estimateur convergent pour le paramètre de dispersion Φ . Cette estimation est moins lourde que celle du maximum de vraisemblance. En effet, si on pose t la taille de l'échantillon observé ou concrètement le nombre d'éléments du triangle supérieur et p le nombre de paramètres explicatifs du modèle, nous aurons : $\widehat{\Phi} = \frac{D}{t-p}$

Plus la valeur de D est faible, plus le modèle est crédible. Cependant, la comparaison avec D nécessite que les modèles aient la même composante aléatoire, ce qui est très limitatif pour nos interprétations.

b) Application numérique :

La table qui suit donne les valeurs de quelques statistiques relatives à la qualité d'ajustement du modèle de poisson.

Critère d'évaluation de l'adéquation

Critère	DDL	Valeur	Valeur/DDL
Deviance	45	3979.7601	88.4391
Scaled Deviance	45	3979.7601	88.4391
Pearson Chi-Square	45	3810.7794	84.6840
Scaled Pearson X2	45	3810.7794	84.6840
Log Likelihood		784785.2772	
Full Log Likelihood		-2275.0065	
AIC (smaller is better)		4592.0129	
AICC (smaller is better)		4613.0129	
BIC (smaller is better)		4637.9957	

Algorithm converged.

D'après le tableau ci-dessus, la valeur estimée du paramètre de dispersion du modèle de poisson est $\widehat{\Phi} = 88,44$.

A l'instar du modèle de poisson, on extrait de la feuille de sortie SAS le tableau relatif au calcul de quelques statistiques d'ajustement :

Critère d'évaluation de l'adéquation

Critère	DDL	Valeur	Valeur/DDL
Deviance	45	11.1505	0.2478
Scaled Deviance	45	67.8041	1.5068
Pearson Chi-Square	45	6.8356	0.1519
Scaled Pearson X2	45	41.5661	0.9237
Log Likelihood		-484.7263	
Full Log Likelihood		-484.7263	
AIC (smaller is better)		1013.4526	
AICC (smaller is better)		1036.9875	
BIC (smaller is better)		1061.6250	

Algorithm converged.

D'après le tableau ci-dessus, la déviance du modèle est de l'ordre de 11,15 et la valeur estimée du paramètre de dispersion du modèle de poisson est $\hat{\Phi} = 0.25$.

Chapitre III. La technique du Bootstrap

I. Idée et principe de base de la méthode du Bootstrap :

I.1 Présentation générale :

La méthode du Bootstrap est une technique statistique de ré échantillonnage utilisant des simulations de Monte-Carlo. Elle a été introduite par Efron en 1979 principalement pour réduire le biais d'une estimation ou bien pour donner la variabilité de l'estimateur d'un paramètre d'intérêt dans un contexte non paramétrique.

Etant une méthode de simulation par excellence, Bootstrap est une bonne alternative pour pallier à la non suffisance des données en cas d'approches asymptotiques. Cependant, la mise en œuvre et l'application de cette technique nécessitent la disponibilité d'outils informatiques robustes et performants.

I.2 L'approche du Bootstrap :

Littéralement, Bootstrap signifie « se hisser en tirant sur ses propres lacets ». Statistiquement, elle désigne le fait de se baser sur échantillon initial pour faire B tirages aléatoires avec remise. Autrement, à partir d'un échantillon de taille T, on obtient un supplément d'information en simulant un nombre B assez important d'échantillons de même taille.

Dans le cadre de notre projet, on utilise la méthode du Bootstrap surtout pour remédier au problème des formules analytiques trop complexes en les remplaçant par des calculs beaucoup plus simples basés sur les échantillons simulés.

L'hypothèse de base de la technique Bootstrap est que les éléments de l'échantillon initial sont indépendants (d'où la remise) et sont issus d'une même loi de probabilité (i.i.d).

II. Application du Bootstrap au triangle des règlements :

Nous souhaitons ici calculer le degré d'imprécision dû à l'estimation de la provision dans le cadre d'un modèle statistique. Le principal objectif de cette section est l'estimation des intervalles de confiance pour chacune des provisions calculées pour le modèle Gamma.

II.1 Les résidus de Pearson :

L'hypothèse de base de la technique de Bootstrap n'étant pas forcément toujours vérifiée, il faut chercher à la contourner en transformant les données du triangle de liquidation. En effet, afin de normaliser nos données, on privilégie l'usage des résidus de Pearson ainsi construits : $r_{ij} = \frac{y_{ij} - \hat{\mu}_{ij}}{\sqrt{V(\hat{\mu}_{ij})}}$ pour les modèles GLM, quand $i+j \leq n$

y_{ij} : est la valeur réellement observée ou bien la valeur figurant dans la cellule (i,j) du triangle de liquidation de base.

$\hat{\mu}_{ij}$: est la valeur prévue par le modèle pris.

II.2 Les étapes du Bootstrap aux résidus de Pearson appliquées au modèle de Gamma

Nous estimons dans un premier temps les incréments de paiements du triangle supérieur par le modèle de Gamma déjà établi dans la partie de la modélisation GLM.

I \ j	0	1	2	3	4	5	6	7	8	9	10
1997	828.8	2978.2	3753.5	2268.2	1291.8	595.8	469.4	223.0	198.1	10.1	4.0
1998	1193.1	4286.8	5402.8	3264.8	1859.5	857.7	675.7	321.0	285.2	14.6	-

I \ j	0	1	2	3	4	5	6	7	8	9	10
1999	614.1	2206.5	2780.9	1680.4	957.1	441.4	347.8	165.2	146.8	-	-
2000	754.8	2712.1	3418.7	2065.6	1176.5	542.6	427.5	203.1	-	-	-
2001	657.3	2362.6	2977.0	1798.9	1024.6	472.6	372.3	-	-	-	-
2002	694.5	2495.9	3145.3	1900.6	1082.5	499.3	-	-	-	-	-
2003	1123.0	4035.5	5085.6	3073.2	1750.4	-	-	-	-	-	-
2004	1320.9	4746.5	5982.1	3614.9	-	-	-	-	-	-	-
2005	993.0	3568.2	4496.9	-	-	-	-	-	-	-	-
2006	875.4	3145.7	-	-	-	-	-	-	-	-	-
2007	1122.0	-	-	-	-	-	-	-	-	-	-

Les résidus de Pearson qui en résultent sont :

I \ j	0	1	2	3	4	5	6	7	8	9	10
1997	-0,02	-0,41	0,17	0,06	-0,27	0,04	0,07	-0,05	0,03	0,01	0,0
1998	-0,21	-0,47	-0,50	-0,83	-0,25	-0,25	-0,07	0,27	0,13	-0,02	-
1999	0,05	0,21	0,16	0,12	0,36	0,08	-0,07	-0,02	-0,08	-	-
2000	0,11	-0,16	-0,07	0,33	0,18	0,11	-0,01	-0,12	-	-	-
2001	0,07	0,09	-0,26	-0,16	-0,13	-0,01	0,07	-	-	-	-
2002	0,05	0,09	-0,01	-0,18	0,13	-0,07	-	-	-	-	-
2003	-0,19	0,86	0,47	0,14	-0,32	-	-	-	-	-	-
2004	-0,10	-0,47	0,25	0,49	-	-	-	-	-	-	-
2005	-0,05	0,33	-0,16	-	-	-	-	-	-	-	-
2006	0,06	-0,22	-	-	-	-	-	-	-	-	-
2007	0,00	-	-	-	-	-	-	-	-	-	-

On doit maintenant effectuer des tirages avec remise pour les éléments du triangle des résidus. Voici l'exemple d'un premier échantillon :

I \ j	0	1	2	3	4	5	6	7	8	9	10
1997	0,06	0,04	-0,04	-0,04	0,00	0,00	0,00	-0,21	-0,47	-0,47	-0,51
1998	-0,83	-0,83	-0,89	-0,25	-0,25	-0,24	-0,07	-0,07	0,27	0,27	-
1999	0,12	0,36	0,02	-0,07	-0,01	-0,08	-0,16	-0,16	-0,07	-	-
2000	-0,07	0,18	0,15	0,11	-0,01	-0,16	-0,16	-0,01	-	-	-
2001	0,18	0,11	0,15	-0,01	-0,16	-0,16	-0,01	-	-	-	-
2002	-0,01	0,06	0,06	0,09	-0,01	0,13	-	-	-	-	-
2003	-0,07	-0,07	-0,11	-0,19	0,86	-	-	-	-	-	-
2004	0,86	0,47	0,47	0,14	-	-	-	-	-	-	-
2005	0,14	0,14	-0,31	-	-	-	-	-	-	-	-
2006	-0,47	-0,47	-	-	-	-	-	-	-	-	-
2007	0,25	-	-	-	-	-	-	-	-	-	-

La construction du triangle de paiements simulé à partir des résidus est l'étape qui suit. La formule utilisée à cette fin est $C_{ij}^b = r_{ij}^{b*} \sqrt{V(\widehat{\mu}_{ij})} + \widehat{\mu}_{ij}$.

I \ j	0	1	2	3	4	5	6	7	8	9	10
1997	927,59	3048,33	3682,50	2197,16	1291,88	595,89	469,48	122,91	555,65	743,65	801,81
1998	140,80	2952,95	4069,00	2867,38	1462,02	461,02	560,97	206,28	715,16	444,57	-
1999	812,51	2788,44	2908,46	1563,62	935,82	301,72	80,71	101,85	33,60	-	-
2000	641,61	3006,63	3595,55	2242,90	1155,96	284,77	169,64	181,51	-	-	-
2001	951,84	2539,34	3154,29	1778,45	766,71	214,71	350,72	-	-	-	-
2002	672,91	2602,27	3233,70	2051,08	1065,19	715,73	-	-	-	-	-
2003	1010,75	3922,89	4778,65	2766,18	3126,18	-	-	-	-	-	-
2004	2696,79	5507,74	6743,45	3839,76	-	-	-	-	-	-	-
2005	1217,82	3792,84	3988,60	-	-	-	-	-	-	-	-
2006	126,6	2397,18	-	-	-	-	-	-	-	-	-
2007	1519,81	-	-	-	-	-	-	-	-	-	-

Et la réserve correspondant à ce nouveau tableau d'incrément simulé est : 55 564,62

Ce traitement sera refait 100 fois pour aboutir à un échantillon de 100 réserves. Sur la base de ces réserves simulées, on calcule la réserve finale retenue comme étant la moyenne de toutes les valeurs de l'échantillon. De même, on substitue aux formules analytiques de variance complexes déjà vues la formule directe du calcul de la variance à partir du même échantillon de réserves simulé :

$$R^B = \frac{1}{100} \sum_{b=1}^{100} R_b$$

$$V(R^B) = \frac{1}{100} \sum_{b=1}^{100} (R_b - R^B)^2$$

44 572,58	48 626,10	43 525,19	43 608,01	38 849,00	43 153,54	45 666,04	32 265,18
41 523,34	40 399,76	39 278,03	39 118,99	36 690,86	44 112,77	33 127,41	43 533,66
35 529,89	39 139,38	45 428,66	35 618,42	35 207,78	36 149,01	37 829,44	39 867,31
38 445,20	33 573,13	45 765,83	35 195,22	43 331,48	45 590,28	48 058,59	54 540,79
43 600,64	33 760,50	50 264,54	48 704,40	40 138,81	40 045,88	40 413,99	39 629,44
42 043,24	47 405,17	49 174,13	43 750,84	46 332,93	35 513,25	33 894,85	40 944,48
33 367,86	45 782,73	54 041,04	45 306,74	42 412,24	51 116,32	33 961,52	37 280,57
43 673,96	45 695,83	45 541,12	51 546,81	39 589,99	50 908,90	38 143,63	41 374,47
34 286,41	35 373,24	45 237,70	39 679,18	30 090,17	44 079,81	50 158,73	42 267,68
47 728,85	42 827,40	38 522,00	49 470,54	43 404,70	44 367,27	41 532,18	47 788,80
46 402,86	49 480,75	45 176,72	42 330,46	40 604,32	40 190,98	50 057,51	38 683,13
47 504,21	44 655,99	42 160,99	47 021,85	39 018,27	44 308,54	48 360,48	36 686,49
41 315,49	62206,5908	37876,0939	54212,2612				

On trouve donc : $R^B=42577,5$ et $\sigma^B=5720,187$. Pour construire un intervalle de confiance de la réserve globale obtenue par la technique du Bootstrap, on suppose que R^B suit la loi normale. Donc, au seuil $\alpha=5\%$, $IC_{5\%}(R^B) = [31 137,13 ; 54 017,88]$

Chapitre IV. La méthode de Munich Chain Ladder

Le modèle de Munich-Chain-Ladder propose d'utiliser non seulement le triangle des paiements cumulés mais aussi celui des charges afin de réduire les écarts entre les IBNR basés sur les sinistres payés et ceux basés sur les sinistres survenus. En effet, cette méthode contrairement aux autres méthodes stochastiques, intègre la corrélation naturelle existante entre les paiements et les charges.

I. Les hypothèses du modèle :

Pour mieux éclaircir les hypothèses de travail, nous allons énoncer et rappeler quelques notations dont on aura besoin.

Pour commencer, on rappelle que $C_{i,j}$ désigne le montant des paiements cumulés de l'année d'origine i après j années de développement. On note par $I_{i,j}$ la charge relative aux sinistres de l'année d'origine i vue à la fin de l'année j . Nous noterons de plus les informations disponibles sur les paiements et les charges au bout de j années de développement par $P_i(j)=(P_{i,k})_{k=0,\dots,j}$, $I_i(j)=(I_{i,k})_{k=0,\dots,j}$ et $B_i(j)=(P_i(j), I_i(j))$. En dernier lieu et pour $i,j=0\dots n$, nous notons respectivement par Q_{ij} et Q_{ij}^{-1} les quotients $\frac{C_{ij}}{I_{ij}}$ et $\frac{I_{ij}}{C_{ij}}$.

Le modèle Munich-Chain-Ladder repose sur plusieurs hypothèses qu'on peut fractionner en deux sous-groupes selon les paiements et les charges.

I.1 Les hypothèses relatives aux paiements :

- La première hypothèse (H_1) suppose que les couples $(C_{1,j}, I_{1,j})_{j=0,\dots,n}$, $(C_{2,j}, I_{2,j})_{j=0,\dots,n}$, ..., $(C_{n,j}, I_{n,j})_{j=0,\dots,n}$ sont indépendants ce qui implique l'indépendance des paiements par année d'origine.
- La deuxième hypothèse (H_2) suppose, pour chaque j appartenant à l'intervalle $[0, n-1]$, l'existence d'un paramètre $f_j^{(p)}$ tel que :

$$E \left[\frac{C_{i,j+1}}{C_{i,j}} / P_i(j) \right] = f_j^{(p)} \text{ pour } i=0\dots n$$

- La troisième hypothèse (H_3) suppose que pour chaque j appartenant à l'intervalle $[0, n-1]$, il existe un paramètre $\sigma_j^{(p)} > 0$ tel que :

$$V \left[\frac{C_{i,j+1}}{C_{i,j}} / P_i(j) \right] = \frac{\sigma_j^{(p)2}}{C_{i,j}} \text{ pour } i=0\dots n$$

- La dernière hypothèse (H_4) dit que les facteurs standardisés $\text{Res}(\frac{C_{i,j+1}}{C_{i,j}} / P_i(j))$ sont linéairement corrélés aux ratios (I/P) standardisés ce qui signifie que pour chaque j appartenant à l'intervalle $[0, n-1]$ et pour chaque i appartenant à l'intervalle $[0, n]$ il existe un paramètre $\lambda^{(p)}$ tel que :

$$E \left[\text{Res} \left(\frac{C_{i,j+1}}{C_{i,j}} / P_i(j) \right) / B_i(j) \right] = \lambda^{(p)} \text{Res} (Q_{ij}^{-1} / P_i(j))$$

$$\text{Avec } \text{Res} \left(\frac{C_{i,j+1}}{C_{i,j}} / P_i(j) \right) = \frac{C_{i,j+1}}{C_{i,j}} f_j^{(p)} - \sqrt{C_{i,j}} \text{ et } \text{Res} (Q_{ij}^{-1} / P_i(j)) = \frac{Q_{ij}^{-1} - E(Q_{ij}^{-1} / P_i(j))}{\sigma((Q_{ij}^{-1} / P_i(j)))}$$

I.2 Les hypothèses relatives aux charges :

Pour les hypothèses concernant les charges, on remarque qu'ils sont analogues à (H_2), (H_3) et (H_4). Ainsi, les hypothèses sur les charges sont :

- (H_5) : pour $j=0,\dots,n-1$, il existe un paramètre $f_j^{(I)}$ tel que :

$$E \left[\frac{I_{i,j+1}}{I_{i,j}} / I_i(j) \right] = f_j^{(I)} \text{ pour } i=0\dots n$$

- (H_6) : pour $j=0,\dots,n-1$, il existe un paramètre $\sigma_j^{(I)} > 0$ tel que :

$$V \left[\frac{I_{i,j+1}}{I_{i,j}} / I_i(j) \right] = \frac{\sigma_j^{(I)2}}{I_{i,j}} \text{ pour } i=0\dots n$$

➤ (H7) pour chaque j appartenant à l'intervalle $[0, n-1]$ et pour chaque i appartenant à l'intervalle $[0, n]$ il existe un paramètre $\lambda^{(I)}$ tel que :

$$E \left[\text{Res} \left(\frac{I_{i,j+1}}{I_{i,j}} / I_i(j) \right) / B_i(j) \right] = \lambda^{(I)} \text{Res} (Q_{ij} / I_i(j))$$

$$\text{Avec } \text{Res} \left(\frac{I_{i,j+1}}{I_{i,j}} / I_i(j) \right) = \frac{I_{i,j+1} - f_j^{(I)}}{\sigma_j^{(I)}} \sqrt{I_{i,j}} \text{ et } \text{Res} (Q_{ij} / I_i(j)) = \frac{Q_{ij} - E((Q_{ij} / I_i(j)))}{\sigma((Q_{ij} / I_i(j)))}$$

Cela signifie que le paramètre $\lambda^{(I)}$ est la pente de la droite de régression des facteurs de développement du triangle des charges standardisées et le triangle des ratios standardisés Q_{ij} .

II. Estimation des paramètres :

Les paramètres $f_j^{(P)}$ et $\sigma_j^{(P)}$ sont estimés de la même manière que dans le modèle de Mack respectivement par $\frac{\sum_{i=0}^{n-j} C_{i,j+1}}{\sum_{i=0}^{n-j} C_{i,j}}$ et $\frac{1}{n-j-1} \sum_{i=0}^{n-j} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - f_j^{(P)} \right)^2$. Quant à $\sigma_{n-1}^{(P)}$ elle est estimée par $\min \left(\left(\frac{\sigma_{n-2}^{(P)}}{\sigma_{n-3}^{(P)}} \right)^2, \sigma_{n-2}^{(P)2}, \sigma_{n-3}^{(P)2} \right)$.

	0	1	2	3	4	5	6	7	8	9
$f_j^{(P)}$	4.7133	1.9971	1.30113	1.12599	1.05346	1.04188	1.02254	1.02025	1.00090	1.00033
$\sigma_j^{(P)}$	34.0742	10.7163	5.81959	4.03191	1.27772	1.37527	2.70895	1.86703	0.01442	0.00414

Par la suite on calcule les estimateurs des résidus standardisés des facteurs des

paiements grâce à la formule suivante : $\text{Res} \left(\frac{C_{i,j+1}}{C_{i,j}} / P_1(j) \right) = \frac{C_{i,j+1} - f_j^{(P)}}{\sigma_j^{(P)}} \sqrt{C_{i,j}}$

I \ J	0	1	2	3	4	5	6	7	8	9
1997	-0.64020	1.54731	0.43155	-0.89386	0.85808	0.91962	-0.37718	0.03634	0.74078	0
1998	0.39112	0.32388	-1.40049	0.20827	-1.40926	0.15211	1.43261	0.90782	-0.67169	.
1999	-0.09264	-0.32160	-0.03918	1.39461	0.39301	-1.43220	-0.33512	-1.08376	.	.
2000	-0.97952	-0.10148	1.23294	0.77632	1.09569	-0.49853	-0.83251	.	.	.
2001	-0.41898	-1.17192	-0.49686	-0.35320	0.10840	0.91184
2002	-0.27425	-0.46381	-0.79044	0.74425	-0.95444
2003	2.44615	-0.42801	-0.52826	-1.38974
2004	-0.25424	1.61300	1.47648
2005	0.73337	-1.00684
2006	-0.79513

On estime les coefficients $E [Q_{ij}^{-1} / P_i(j)]$ de triangle I/P par les rapports $\frac{\sum_{k=0}^{n-j} I_{k,j}}{\sum_{k=0}^{n-j} C_{k,j}}$ et

on les note \hat{q}_j^{-1} , tandis que la variance $V [Q_{ij}^{-1} / P_i(j)]$ est estimée par $\frac{\hat{\rho}_j^{(p)}}{\sqrt{C_{i,j}}}$ où

$$\hat{\rho}_j^{(p)2} = \frac{1}{n-j} \sum_{k=0}^{n-j} C_{k,j} (Q_{ij}^{-1} - \hat{q}_j^{-1})^2.$$

➤ les estimateurs des \hat{q}_j^{-1} :

	0	1	2	3	4	5	6	7	8	9	10
\hat{q}_j^{-1}	12.276	3.7746	1.9028	1.5033	1.3704	1.2199	1.1309	1.1053	1.09	1.0768	1.0232

➤ les estimateurs des $\hat{\rho}_j^{(P)}$

	0	1	2	3	4	5	6	7	8	9
$\hat{\rho}_j^{(P)}$	117.613	25.157	12.183	10.084	13.902	5.27301	2.79279	2.18609	1.0728	1.56299

Par la suite on en déduit les résidus standardisés du triangle des Q_{ij}^{-1} selon la formule suivante $\text{Res}(\widehat{Q}_{ij}^{-1} \setminus P_{i,j}) = \frac{Q_{ij}^{-1} - \widehat{q}_j^{-1}}{\widehat{\rho}_j^{(P)}} \sqrt{C_{i,j}}$

I \ J	0	1	2	3	4	5	6	7	8	9
1997	-0.59626	0.96586	1.38506	-0.83089	-0.59114	-0.33753	-0.07552	0.24023	-1.01492	-0.74078
1998	-0.09926	-0.04313	-1.20388	-0.64016	-0.82891	-0.97057	0.28992	1.16721	0.98218	0.67175
1999	0.30696	0.62460	0.98061	1.02338	-0.00100	1.01108	1.42785	-0.30979	-0.07252	.
2000	-0.74836	-0.47372	-0.04775	-0.36316	-0.75731	-0.74730	-1.34147	-1.21817	.	.
2001	-0.90824	-1.73871	-1.35017	-0.33452	-0.50506	-0.28933	-0.26819	.	.	.
2002	-0.91675	-1.18466	-0.69011	-0.10386	0.15898	1.50984
2003	0.88842	-0.57654	-0.43301	1.94829	2.02721
2004	-0.73297	0.13325	0.26434	-0.89560
2005	-0.09399	0.85250	1.05443
2006	0.19677	1.39545
2007	2.43740

Pour finaliser cette première étape, on calcule $\widehat{\lambda}^{(p)}$ qui vaut $\frac{\sum_{i,j} \text{Res} \left(\frac{C_{i,j+1}}{C_{i,j}} / P_{i(j)} \right) \text{Res}(\widehat{Q}_{ij}^{-1} \setminus P_{i,j})}{\sum_{i,j} \text{Res}(\widehat{Q}_{ij}^{-1} \setminus P_{i,j})^2}$. A la base des tableaux précédents on trouve $\widehat{\lambda}^{(p)}=0,35655$.

Par la suite et afin de compléter la mise en œuvre de ce modèle, on calcule le triangle des facteurs individuels des charges et on en déduit ensuite les paramètres $f_j^{(I)}$ et $\sigma_j^{(I)}$ afin de standardiser les facteurs de ce triangle. Par analogie, on estime $f_j^{(I)}$ et $\sigma_j^{(I)}$ par $\frac{\sum_{i=0}^{n-j} I_{i,j+1}}{\sum_{i=0}^{n-j} I_{i,j}}$ et $\frac{1}{n-j-1} \sum_{i=0}^{n-j} C_{ij} \left(\frac{C_{i,j+1}}{C_{ij}} - f_j^{(I)} \right)^2$. On remarque, pour le cas de $j=n-1$, que la formule de calcul précédente n'aboutit pas à une valeur valable donc comme pour le modèle de Mack on estimera $\sigma_{n-1}^{(I)}$ par $\min \left(\left(\frac{\sigma_{n-2}^{(I)}}{\sigma_{n-3}^{(I)}} \right)^2, \sigma_{n-2}^{(I)2}, \sigma_{n-3}^{(I)2} \right)$. Tandis que les résidus seront estimés

$$\text{par } \text{Res} \left(\frac{I_{i,j+1}}{I_{i,j}} / P_{i(j)} \right) = \frac{I_{i,j+1} - f_j^{(I)} I_{i,j}}{\sigma_j^{(I)}} \sqrt{I_{i,j}}$$

I \ J	0	1	2	3	4	5	6	7	8	9
1997	1.69441	1.14562	0.89729	0.99089	0.99267	0.98908	0.99619	0.99320	0.98822	0.95989
1998	1.63661	0.95328	1.00204	0.99657	0.96190	1.01101	1.04574	1.00872	0.98880	.
1999	1.36500	0.98186	1.02265	1.00658	0.98303	0.93995	0.95213	0.99198	.	.
2000	1.37570	1.05067	1.07357	1.00316	0.99487	0.95970	0.98756	.	.	.

➤ Estimation des \hat{q}_j

	1	2	3	4	5	6	7	8	9	10	11
\hat{q}_j	0.081460	0.26493	0.52554	0.66519	0.72972	0.81973	0.88423	0.90471	0.91744	0.92866	0.97736

➤ Estimation de $\hat{\rho}_j^{(1)}$

	1	2	3	4	5	6	7	8	9	10
$\hat{\rho}_j^{(1)}$	2.52495	3.53924	4.66062	5.34058	8.35198	3.86365	2.31885	1.88367	0.94314	1.40002

➤ Triangles des résidus des Q_{ij}

I \ J	1	2	3	4	5	6	7	8	9	10	11
1997	0.58478	-0.96674	-1.32057	0.94714	0.55661	0.36834	0.06852	-0.32231	0.99765	0.70711	0
1998	-0.20331	-0.07355	1.08070	0.63996	0.80009	0.96523	-0.26932	-1.16056	-1.00233	-0.70711	.
1999	-0.67886	-0.67379	-1.03799	-1.32866	-0.23500	-0.98799	-1.43057	0.26723	0.00468	.	.
2000	0.77044	0.36661	-0.02936	0.34601	0.77761	0.78486	1.33796	1.21564	.	.	.
2001	1.28626	1.98931	1.58913	0.36086	0.52355	0.35717	0.29341
2002	1.30825	1.18260	0.67698	0.03068	-0.44301	-1.48761
2003	-1.14278	0.31473	0.26032	-1.75918	-1.97985
2004	0.60482	-0.21105	-0.28157	0.76318
2005	-0.21360	-0.72607	-0.93765
2006	-0.52674	-1.20205
2007	-1.78926

On estime $\hat{\lambda}^{(1)}$ par $\frac{\sum_{i,j} \text{Res} \left(\frac{I_{i,j+1}}{I_{i,j}} / I_{i,j} \right) \text{Res}(\widehat{Q}_{i,j} \setminus I_{i,j})}{\sum_{i,j} \text{Res}(\widehat{Q}_{i,j} \setminus I_{i,j})^2}$ et on a $\lambda^{(1)} = 0.2785944$.

Après avoir évalué les estimateurs nécessaires on passe maintenant à la dernière étape qui consiste à compléter les deux rectangles. Pour cela on utilisera les formules récursives ci-dessous:

➤ Pour les paiements : $C_{i,j+1} = \widehat{C}_{i,j} [f_j^{(p)} + \lambda^{(p)} \frac{\sigma_j^{(p)}}{\rho_j^{(p)}} (\widehat{Q}_{i,j}^{-1} - \widehat{q}_j^{-1})]$

Avec $\widehat{Q}_{i,j}^{-1} = \frac{\widehat{I}_{i,j}}{\widehat{C}_{i,j}}$ et $C_{i,n-i+1} = C_{i,n-i+1}$.

➤ Pour les charges $\hat{I}_{i,j+1} = \hat{I}_{i,j} [f_j^{(1)} + \hat{\lambda}^{(1)} \frac{\hat{\sigma}_j^{(1)}}{\hat{\rho}_j^{(1)}} (\widehat{Q}_{i,j} - \hat{q}_j)]$

Avec $\hat{I}_{i,n-i+1} = I_{i,n-i+1}$ et $\widehat{Q}_{i,j} = \frac{\widehat{C}_{i,j}}{\widehat{I}_{i,j}}$.

➤ Rectangle des paiements

I \ J	0	1	2	3	4	5	6	7	8	9	10
1997	792	3119	7155	9522	10370	11036	11631	11783	12029	12041	12045
1998	847	4380	8977	10908	12370	12831	13392	14143	14631	14643	14647,95
1999	708	3253	6300	8179	9718	10287	10518	10662	10669	10678,56	10682,08
2000	933	3378	6683	9282	10753	11473	11880	11902	12054,55	12064,23	12067,98
2001	783	3291	5852	7393	8202	8653	9132	9313,09	9494,60	9502,74	9505,79
2002	783	3429	6557	8159	9458	9845	10330,77	10798,03	11039,31	11049,02	11052,62
2003	816	6227	12074	15372	16614	17621,22	18754,87	20266,03	20794,84	20813,75	20820,64
2004	1156	5154	11534	15930	17774,52	18690,47	19482,02	20035,29	20445,41	20463,08	20469,69

2005	903	5007	9236	12227,51	13879,95	14632,79	15348,88	16028,24	16384,69	16399,08	16404,42
2006	977	3758	7831,96	10421,47	11850,94	12497,02	13117,01	13718,74	14026,25	14038,59	14043,17
2007	1122	6280,22	14324,51	19755,06	22740,97	24023,80	25325,73	26763,64	27395,18	27419,55	27428,54

➤ Rectangle des charges

I \ J	0	1	2	3	4	5	6	7	8	9	10
1997	7749	13130	15042	13497	13374	13276	13131	13081	12992	12839	12324
1998	10058	16461	15692	15724	15670	15073	15239	15936	16075	15895	15256,30
1999	9652	13175	12936	13229	13316	13090	12304	11715	11621	11487,93	11027,25
2000	8765	12058	12669	13601	13644	13574	13027	12865	12894,79	12750,27	12241,04
2001	6623	9913	9877	10824	10604	10414	10256	10270,72	10263,13	10146,65	9740,45
2002	6595	11198	11796	12171	13176	12800	12361,05	12009,57	11985,44	11848,72	11373,91
2003	13002	22360	22395	25545	26400	25751,13	24357,39	22789,34	22706,28	22445,53	21544,95
2004	11260	19695	22293	22808	23311,80	22895,22	22362,67	22159,85	22133,77	21882,13	21005,82
2005	10753	20417	18809	19166,80	19351,02	18961,15	18322,36	17821,01	17786,04	17583,19	16878,61
2006	12717	16337	16259,50	16483,79	16622,75	16284,21	15719,27	15261,15	15230,02	15056,26	14452,90
2007	23376	35439,86	33815,22	33184,36	33211,41	32487,85	31148,46	29876,57	29800,10	29459,39	28278,38

III. Le calcul des provisions :

Après avoir complété les deux triangles inférieurs on calcule la provision par exercice à partir du rectangle des paiements de la même manière que pour le modèle de Mack tandis que pour le rectangle des charges, la provision par exercice est obtenue en faisant la différence entre la charge ultime et le paiement cumulé.

Donc les provisions par exercice données par le rectangle des paiements sont:

Année d'origine	Paiement ultime	Elément diagonale	Provision par exercice
1997	12045	12045	0
1998	14647,95	14643	4,95
1999	10682,08	10669	13,08
2000	12067,98	11902	165,98
2001	9505,79	9132	373,79
2002	11052,62	9845	1207,62
2003	20820,64	16614	4206,64
2004	20469,69	15930	4539,69
2005	16404,42	9236	7168,42
2006	14043,17	3758	10285,17
2007	27428,54	1122	26306,54
			54271,89

D'où la provision à la fin de l'année 2007 est égale à 54271,89031
Le tableau des provisions obtenues suite au triangle des charges :

Année d'origine	Charge ultime	Paiement cumulé	Provision par exercice
1997	12324	12045	279
1998	15256,30	14643	613,30

1999	11027,25	10669	358,25
2000	12241,04	11902	339,04
2001	9740,45	9132	608,45
2002	11373,91	9845	1528,91
2003	21544,95	16614	4930,95
2004	21005,82	15930	5075,82
2005	16878,61	9236	7642,61
2006	14452,90	3758	10694,90
2007	28278,38	1122	27156,38
			59227,62

D'où la provision globale est égale à 59227,62.

Si on prend comme base l'évolution de la charge, on observe un écart relatif entre les deux provisions de l'ordre de 8,4%. Dans notre cas, une convergence améliorée des deux résultats prouve que le triangle des charges est correctement complété.

IV. Le calcul du risque d'estimation des provisions :

Les valeurs des MSEP, Se et SeR sont obtenues de la même manière que le modèle de Mack. Elles sont regroupées dans le tableau ci-dessous:

	MSEP(\widehat{R}_t)	Se(\widehat{R}_t)	SeR(\widehat{R}_t)
1997	-	-	-
1998	0,55495825	0,7449552	0,15314494
1999	3,45083792	1,85764311	0,14124506
2000	54310,5419	233,046223	0,91028433
2001	123584,753	351,546232	0,86434466
2002	175683,933	419,146673	0,48227627
2003	443339,894	665,837738	0,27364689
2004	823986,081	907,736791	0,19588693
2005	1310882,6	1144,93781	0,18240283
2006	2766453,08	1663,26579	0,18798806
2007	46116677,8	6790,92614	0,40867673

➤ De même pour les intervalles de confiance au seuil 5% :

	\widehat{R}_t	borne inf	borne sup
1997	-	-	-
1998	4,95211069	3,4919985	6,41222888
1999	13,0772491	9,43626858	16,71822957
2000	165,980609	-290,789987	622,751206
2001	373,793209	-315,237406	1062,823824
2002	1207,62275	386,095273	2029,150229
2003	4206,64492	2901,60295	5511,686883
2004	4539,6897	2760,52559	6318,853809
2005	7168,42444	4924,34633	9412,502552
2006	10285,1689	7025,16798	13545,16987
2007	26306,5364	12996,3212	39616,75165

➤ Pour la réserve totale, les risques d'estimations relatifs sont résumés dans le tableau suivant :

\hat{R}	$\widehat{MSEP}(\hat{R})$	$Se(\hat{R})$	$SeR(\hat{R})$	borne inf	borne sup
54271,8903	11445222,34	3383,07883	0,062335747	47505,73266	61038,048

V. La vérification des hypothèses :

Les hypothèses relatives au triangle des paiements sont déjà vérifiées dans la section du modèle de MACK. Reste à valider les hypothèses correspondantes au triangle des charges.

V.1 L'hypothèse de l'indépendance des charges :

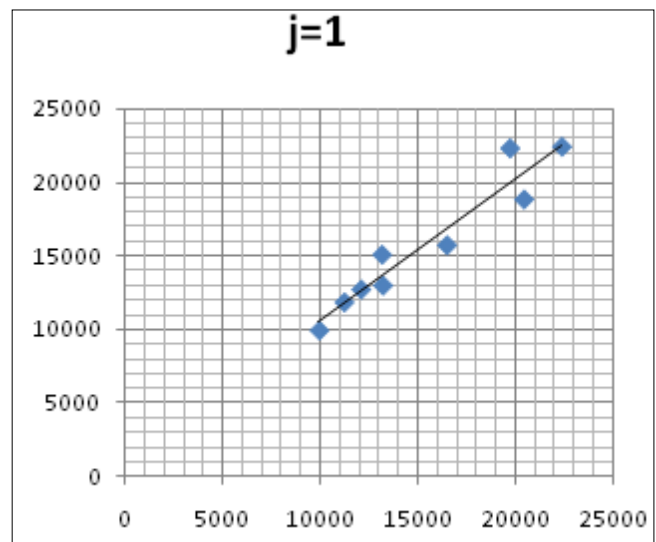
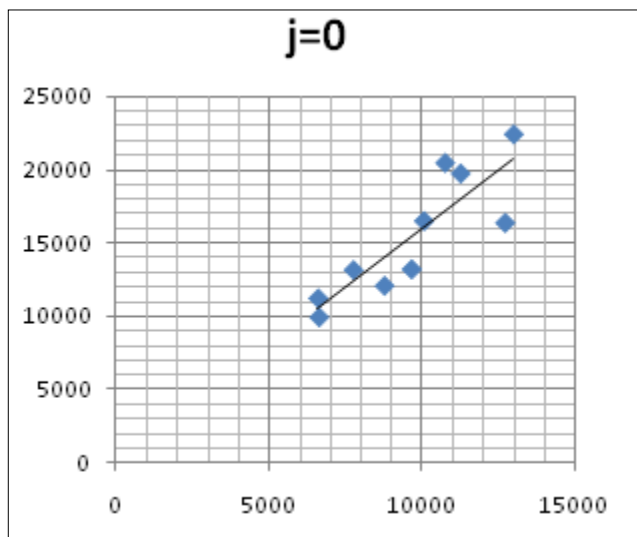
Pour tester l'indépendance des charges par année d'origine on a recours au test non paramétrique de la médiane (Voir le paragraphe de validation du modèle de Mack).

DIAG	Gj	Pj	Zj	E(Zj)	var(Zj)
2	1	1	1	0,50	0,25
3	0	3	0	0,75	0,19
4	2	3	2	1,56	0,37
5	2	3	2	1,56	0,37
6	4	2	2	2,06	0,62
7	6	0	0	2,06	0,62
8	4	2	2	2,06	0,62
9	6	3	3	3,27	0,74
10	1	7	1	2,91	0,80
total			13	16,74	4,58

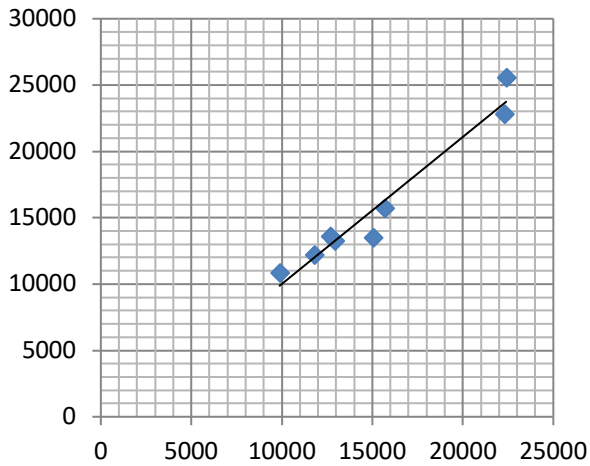
D'après le tableau ci-dessus, on constate que l'hypothèse d'indépendance est vérifiée. En effet, $Z=13$ appartient à l'intervalle $[12.4569, 21.0197]$.

V.2 L'hypothèse de l'existence pour chaque $j=0 \dots (n-1)$ du paramètre $f_j^{(l)}$:

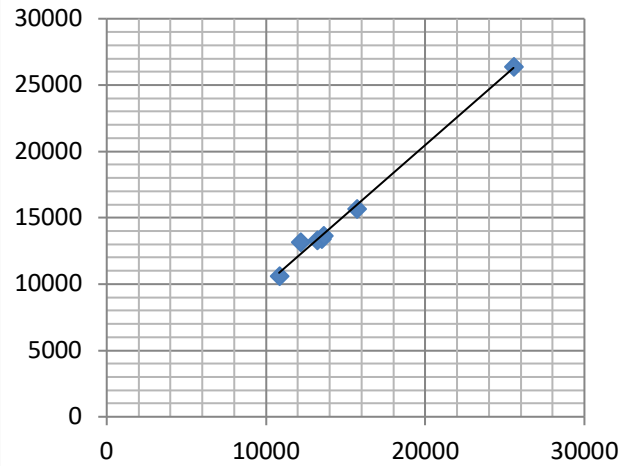
On utilise les C-C plot pour chaque j :



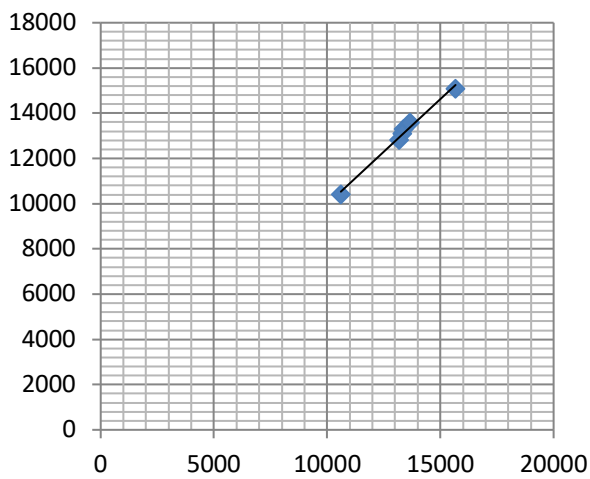
j=2



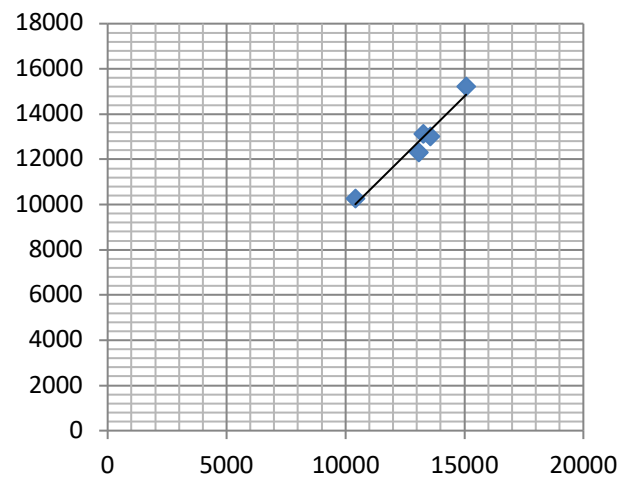
j=3



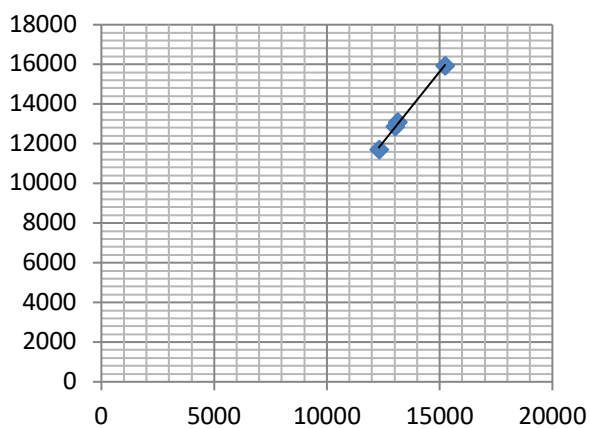
j=4



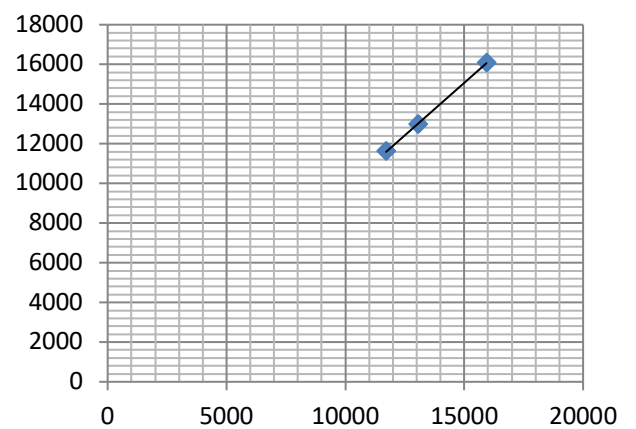
j=5



j=6



j=7



On remarque pour chaque j que les points sont alignés cela justifie l'existence des paramètres $f_j^{(1)}$.

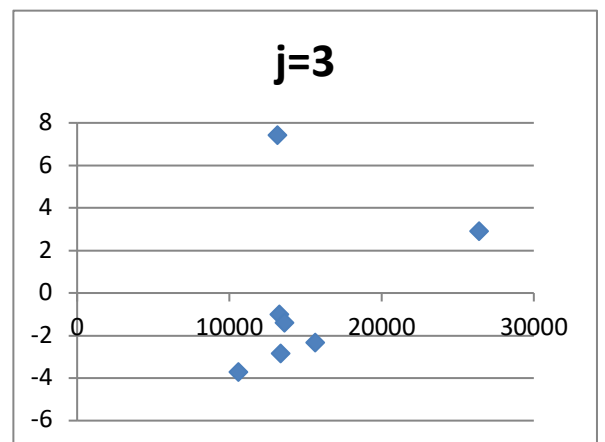
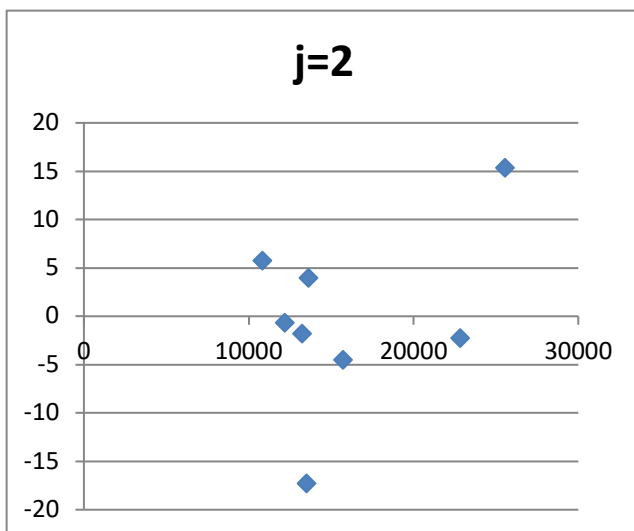
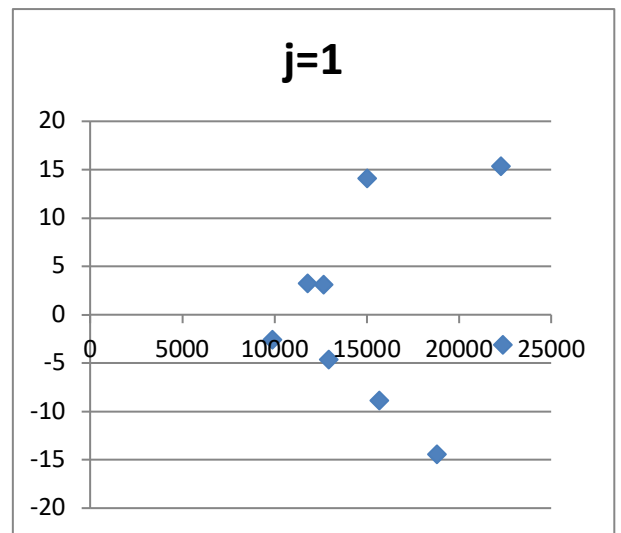
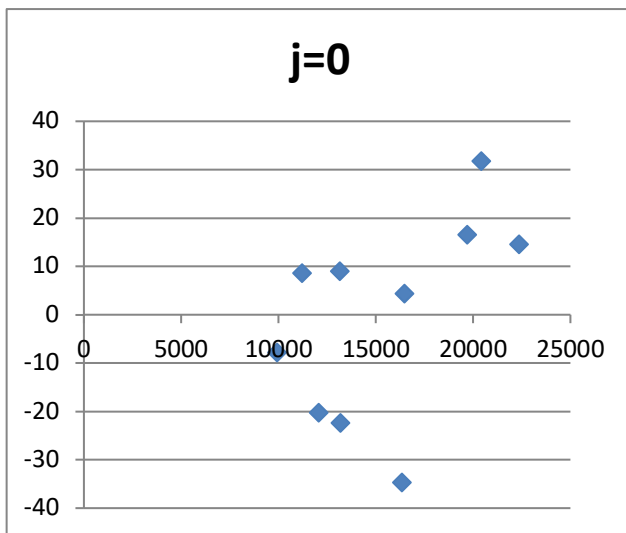
V.3 L'hypothèse de l'existence pour chaque $j=0... (n-1)$ du paramètre $\sigma_j^{(l)}$:

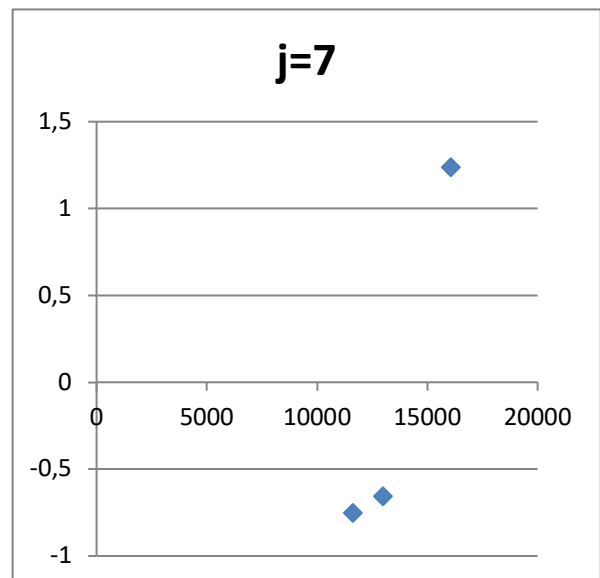
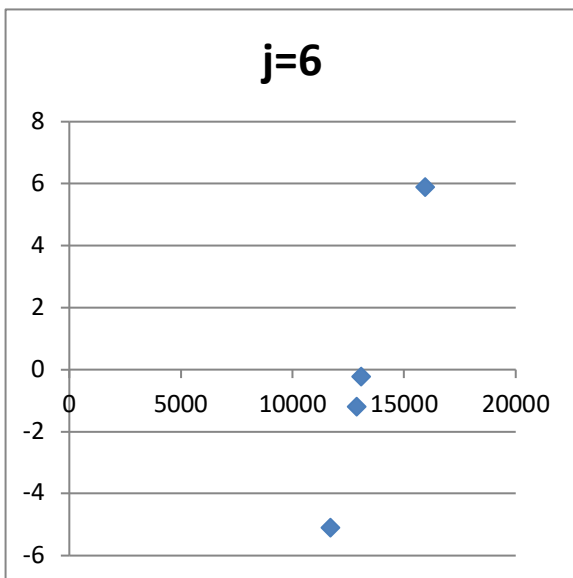
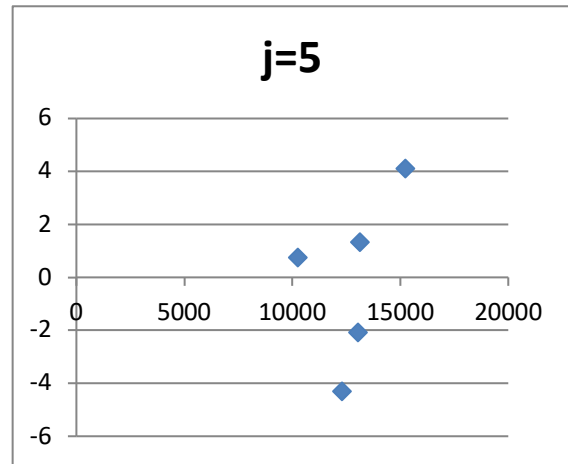
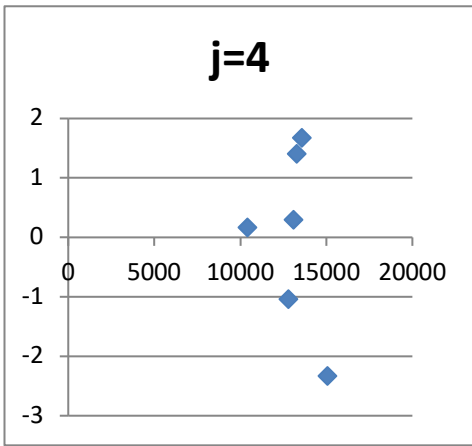
On calcule tout d'abord les résidus standardisés du triangle des charges par la formule

suivante :
$$\frac{c_{i,j+1} - \hat{f}_j c_{ij}}{\sqrt{c_{ij}}}$$

	1	2	3	4	5	6	7	8	9	10
1998	8,9800	14,1182	-17,2946	-2,8281	1,4100	1,3306	-0,2140	-0,6546	-0,0361	-0,0002
1999	4,4336	-8,8690	-4,5423	-2,3404	-2,3256	4,1108	5,8857	1,2374	0,0333	
2000	-22,3406	-4,6545	-1,7800	-0,9953	0,2940	-4,2991	-5,0948	-0,7516		
2001	-20,2879	3,1034	3,9694	-1,4075	1,6808	-2,0771	-1,1979			
2002	-7,7839	-2,5928	5,7224	-3,6991	0,1650	0,7448				
2003	8,5719	3,2796	-0,7070	7,4295	-1,0350					
2004	14,5196	-3,1170	15,3176	2,9153						
2005	16,6292	15,3673	-2,2693							
2006	31,7650	-14,4557								
2007	-34,7039									

Par la suite on trace les résidus en fonction des charges :





D'après les graphes ci-dessus on peut valider l'hypothèse (H_6) puisque les nuages des points ont une structure parfaitement aléatoire.

Conclusion

Conclusion générale

Bibliographie

- *Arrêté du ministre des finances et de la privatisation n° 1548-05 du 6 ramadan 1426 (10 octobre 2005) relatif aux entreprises d'assurances et de réassurance.*
- *Dahir n° 1-60-223 du 12 ramadan 1382 (6 février 1963) portant modification en la forme du dahir du 25 hija 1345 (25 juin 1927) relatif à la réparation des accidents du travail.*
- ***RAPPORT D'ACTIVITE DES ENTREPRISES D'ASSURANCES ET DE REASSURANCE AU MAROC 2009 et 2010.***
- *Provisionnement technique en assurance non-vie, perspectives actuarielles modernes, christian PARTRAT, ECONMICA.*
- *Statistiques de l'assurance 2 partie 1 Assurance non-vie tarification & provisionnement, Université de Rennes 1, Université de Montréal Arthur Charpentier.*
- *NON-Life Insurance Mathematics , Erwin Straub, Springer*
- *Cours régression insea 2ème année, Chaoubi.*
- *Article l'économiste, MAMDA-MCMA S'ALLIE À UNE MUTUELLE DE TAILLE, Édition N° 2486 du 16/03/2007*
- *Article l'économiste, ACCIDENT DU TRAVAIL: QUE DIT LA LOI?, Édition N° 3744 du 2012/03/20*

Annexes

Annexe 1: Le code SAS pour la méthode Chain Ladder Standard et pour le modèle de Mack

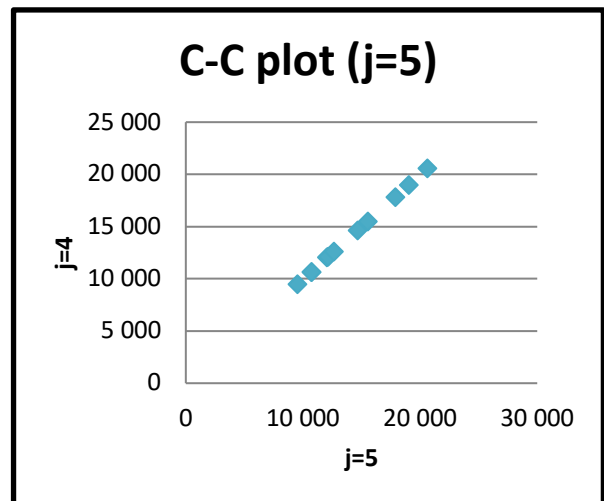
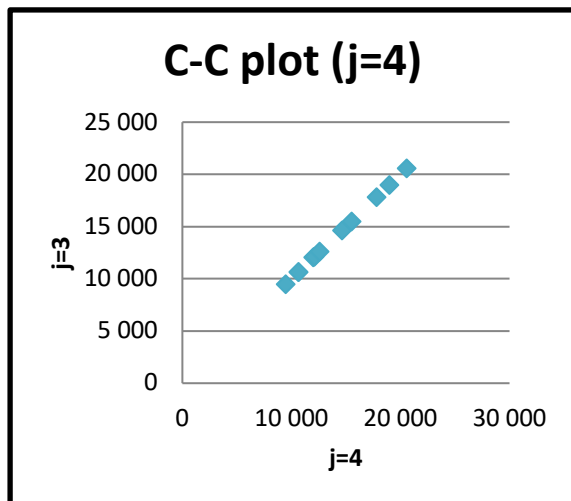
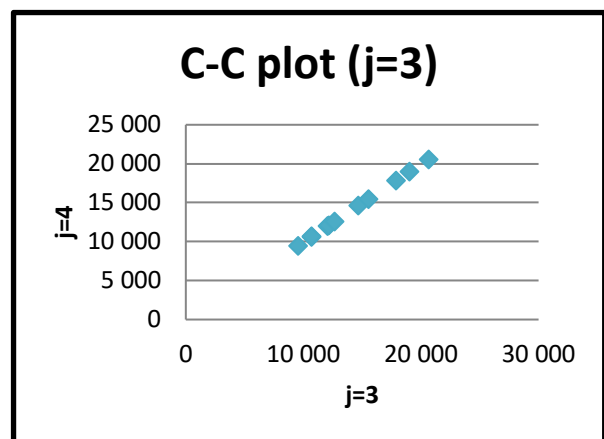
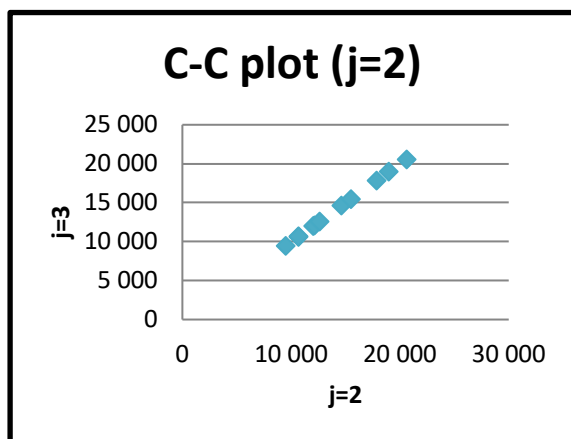
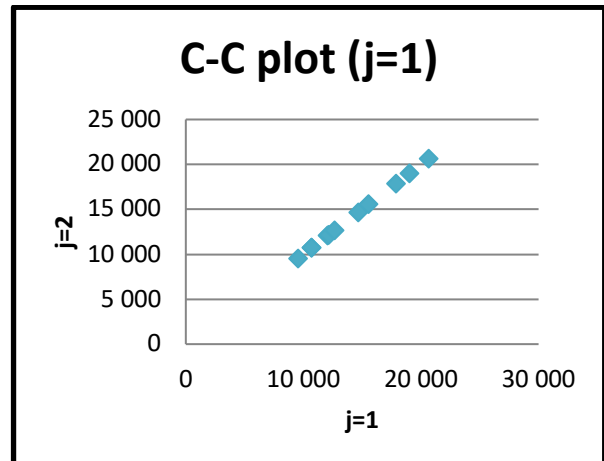
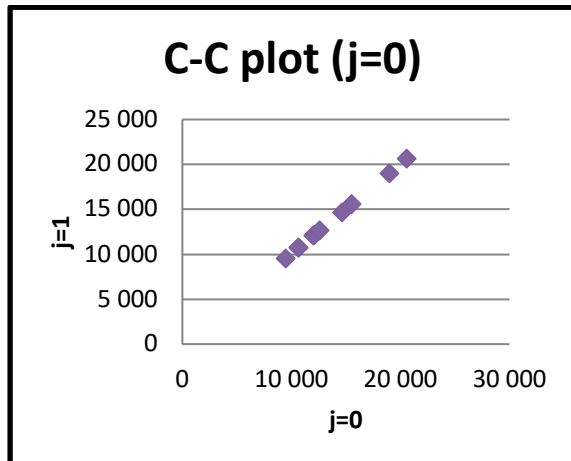
```
/*Construction du Tableau cumulé*/
data sasuser.cumul( drop=i);
set sasuser.Ss;
array cumul [*] _numeric_;
do i=3 to 12;
if cumul(i)>0 then cumul(i)=cumul(i-1) + cumul(i);
end;
run;
/*Calcul des coefficients de passage*/
%MACRO ext(table,var2);
proc sql;
create table base&var2
as SELECT *
FROM &table
where &var2>0;
run;
%MEND;
%macro coef(var1,var2);
proc sql;
create table cof&var1 as
SELECT sum(&var2) / sum(&var1)AS coef&var1
FROM work.base&var2;
run;
%mend;
%ext (sasuser.cumul,F3); %coef (F2,F3);%ext(sasuser.cumul,F4); %coef (F3,F4);
%ext (sasuser.cumul,F5); %coef (F4,F5);%ext(sasuser.cumul,F6);%coef (F5,F6);
%ext (sasuser.cumul,F7); %coef (F6,F7);%ext(sasuser.cumul,F8); %coef (F7,F8);
%ext (sasuser.cumul,F9); %coef (F8,F9);%ext(sasuser.cumul,F10); %coef (F9,F10);
%ext (sasuser.cumul,F11); %coef (F10,F11);%ext(sasuser.cumul,F12); %coef (F11,F12);
proc sql;
create table sasuser.coefficient as select a.*, b.*,c.*,d.*,e.*,f.*,g.*,h.*,i.*, j.*
from coff2 as a ,
    coff3 as b ,
    coff4 as c ,
    coff5 as d,
    coff6 as e ,
    coff7 as f ,
    coff8 as g ,
    coff9 as h ,
    coff10 as i ,
    coff11 as j ;
quit ;
/*tableau avec une seule colonne (transformation du tableau des coefficients)*/
proc iml;
use sasuser.coefficient;
read all var _all_ into x;
t=t(x);
create sasuser.t_coefficient from t;
append from t;
quit;
/*Fusion de tableau cumulé et tableau-coefficient*/
data work.prov;
input COL1 ;
cards ;
0
;
run;
data coef;
set work.prov sasuser.t_coefficient ;
run;
data tab;
merge sasuser.cumul work.coef ;
run;
data sasuser.final;
set work.tab ;
rename COL1 =coef;
run;
/*Compléter le tableau inférieur*/
DATA final1(drop=i);
set sasuser.final;
array c [*] _numeric_;
```

```

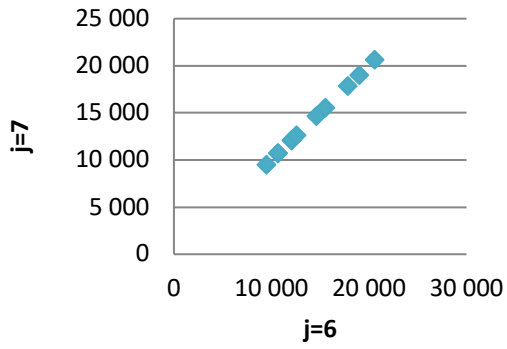
do i=1 to 12;
if c(i)=. then c(i)=0;
end;
run;
proc iml;
USE work.final1;
READ ALL var _num_ INTO mat;
close mat;
print mat;
matrice=J(11,13,0);
print matrice;
DO j=1 to ncol(mat);
do i=1 to nrow(mat);
IF mat[i,j]>0 then matrice[i,j]=mat[i,j];
else matrice[i,j]=matrice[i,j-1]*mat[j-1,13];
end;
END;
print matrice;
create sasuser.tab_comp from matrice;
append from matrice;
quit;
/* Le calcul de la provision*/
data final(drop=col13 );
merge sasuser.tab_comp work.DIAG;
rUN;
data sasuser.provision (drop=i);
set work.final;
array c [*] _numeric_;
do i=0 to 11;
if c(13)= 0 then c(13) = c(12) ;
if c(13)= 1 then c(13) = c(11) ;
if c(13)= 2 then c(13) = c(10) ;
if c(13)= 3 then c(13) = c(9) ;
if c(13)= 4 then c(13) = c(8) ;
if c(13)= 5 then c(13) = c(7) ;
if c(13)= 6 then c(13) = c(6) ;
if c(13)= 7 then c(13) = c(5) ;
if c(13)= 8 then c(13) = c(4) ;
if c(13)= 9 then c(13) = c(3) ;
if c(13)= 10 then c(13) = c(2) ;
END;
PROV= col12 - DIAG ;
RUN;
/* Le calcul de la Provision globale (somme des provisions)*/
proc sql;
create table sasuser.provision_globale as select sum(prov) as total_provision
from sasuser.provision ;
quit;

```

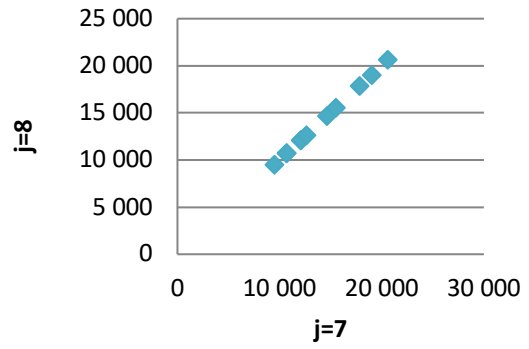
Annexe 2: C-C Plot pour la vérification de l'hypothèse d'indépendance des coefficients de passage pour la méthode Chain Ladder Standard et pour le modèle de Mack



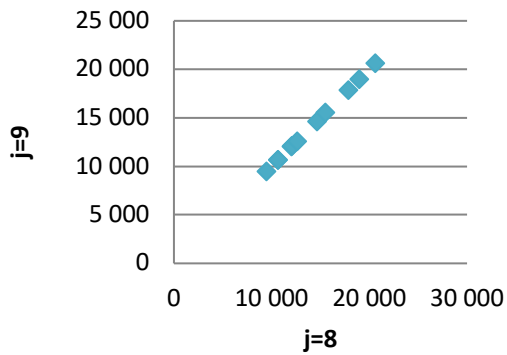
C-C plot (j=6)



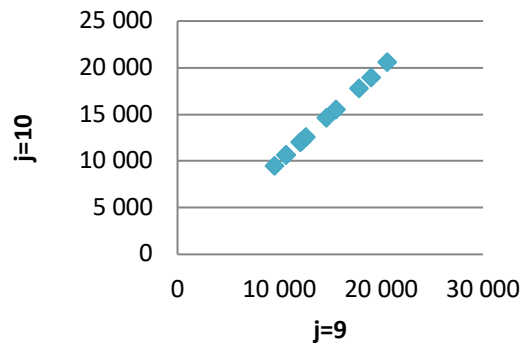
C-C plot (j=7)



C-C plot (j=8)



C-C plot (j=9)



Annexe 3 : Quelques rappels sur les modèles linéaires

□ Fomalisme mathématique :

Supposons que l'on dispose de p variables explicatives et que notre échantillon observé est de taille n . Le modèle linéaire correspondant a la forme générale suivante:

$$Y = X \beta + \varepsilon$$

Avec Y : vecteur colonne de taille n des variables dépendantes.

X : matrice à n lignes et p colonnes des variables explicatives, connues par expérience et qui sont donc non aléatoires,

β : vecteur colonne de taille n contenant les coefficients de régression. Ce sont des paramètres inconnus à estimer,

ε : vecteur des erreurs théoriques aléatoires non observables. Elles représentent l'effet de tous les facteurs qui ont été écartés de l'analyse pour raison de difficulté d'identification. Ces erreurs sont supposées normales, indépendantes, à variance constante σ^2 et de moyenne nulle. La dernière hypothèse se veut d'éviter de biaiser les estimations.

A partir d'un échantillon empirique, et si on note \hat{Y} la valeur prévue de Y et $\hat{\beta}$ le vecteur colonne des estimations des paramètres de régression, on obtient :

$$\hat{Y} = X * \hat{\beta}$$

□ Problèmes d'inadéquation des modèles linéaires :

Les hypothèses de base d'un modèle linéaire conduisent à dire que la variable Y suit une loi normale $N(X\beta, \sigma^2)$. Ce résultat ne semble pas très adéquat pour plusieurs jeux de données. En voici quelques raisons :

1. La loi normale, étant continue, peut ne pas convenir à certains types de données tels que les variables de comptage qui doivent être impérativement discrètes.
2. Il paraît irréaliste de supposer une variance constante pour toutes les observations.

Les GLM, plus souples et moins lourds en hypothèses, viennent permettre l'analyse statistique d'un ensemble plus large de données.

Annexe 4 : A propos du Quantile-Quantile Plot

Le Q-Q Plot est un outil graphique qui, à travers un tracé, permet de juger du degré d'ajustement d'une distribution à un jeu de données. Il autorise également la comparaison de deux distributions jugées similaires.

Comme son nom l'indique, le principe du diagramme Quantile-Quantile repose sur la comparaison des positions des quantiles théoriques et des quantiles empiriques. Concrètement, il s'agit de dessiner le nuage de points $P_i(x_i^*, x_i)$ avec x_i^* est le quantile issu de la distribution théorique et x_i le quantile pris de l'échantillon empirique dont on veut chercher le modèle.

Si la loi sélectionnée s'ajuste bien aux données, le nuage des points P_i s'aligne de manière assez claire autour de la première bissectrice. A la rigueur, le nuage de points pourrait s'ajuster autour d'une droite d'équation : $x_i = ax_i^* + b$. Cela voudrait dire que la distribution choisie est adéquate aux données à une transformation près ; cela ne remet pas en question la forme de la loi mais plutôt ses paramètres.

Pour générer les Q-Q Plot utilisés avec le logiciel SAS, nous avons appliqué le code suivant :

```
/* Générer le Q-Q plot pour la loi normale*/
proc univariate data=sasuser.vect_pnc;
qqplot paiement_sinistre/ normal(mu=est sigma=est);
run;

/* Générer le Q-Q plot pour la loi Gamma*/
proc univariate data=sasuser.vect_pnc;
qqplot paiement_sinistre/ gamma(alpha=est sigma=est theta=est);
run;
```

Annexe 5 : Le principe et les étapes du calcul de la distance K-S

La distance de Kolmogorov-Smirnov est une mesure descriptive d'ajustement. Elle permet d'évaluer la distance entre les lois respectives de deux variables aléatoires ou bien entre une loi théorique prédéfinie et celle d'une variable aléatoire.

La distance de K-S se formule ainsi : $D_{X-Y} = \sup_x |F_X(x) - F_Y(x)|$, avec X et Y deux variables aléatoires .

Sinon, s'il s'agit de la comparaison d'une loi de probabilité à une distribution empirique la distance K-S s'écrit : $D_X = \sup_x |F_X(x) - F^*(x)|$, F^* étant la fonction de répartition de la loi théorique.

Pour notre cas où l'on veut calculer la distance entre la distribution d'un vecteur et celle de deux lois théoriques, il faut commencer inévitablement par construire la fonction de répartition empirique du vecteur aléatoire. Pour ce faire, on segmente les observations par tranches égales, on compte le nombre cumulé d'observations appartenant à chaque intervalle puis on le divise par la taille de l'échantillon. Le résultat est la fonction de répartition empirique.

L'étape suivante consiste à calculer, pour chaque quantile, la valeur de la fonction de répartition de la loi théorique sélectionnée, à savoir la loi Gamma ou la loi normale.

Concernant la loi normale, on s'est basé sur le résultat suivant : Si X suit une loi normale $N(\mu, \sigma^2)$, et si on note Φ la fonction de répartition de la loi normale centrée réduite $N(0,1)$, alors nous avons :

$$F_X(x) = P(X < x) = P\left(\frac{X - \mu}{\sigma} < \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

Par la suite, on utilise la table statistique de la fonction de répartition de la loi $N(0,1)$.

Le cas de la loi gamma est plus délicat vu l'indisponibilité d'une formule de calcul explicite pour la fonction de répartition ou même d'une table statistique servant de base pour le calcul. Les valeurs de cette fonction de répartition furent générées pour notre étude, après avoir estimé les paramètres de la loi Gamma grâce à la moyenne et de l'écart-type, par le logiciel R. Le code est le résultat pour les deux premiers quantiles sont comme suit :

```
> pgamma(500, shape=1.238, rate=.0007)
[1] 0.2005618
> pgamma(1000, shape=1.238, rate=.0007)
[1] 0.3953551
> |
```

Les deux tableaux qui suivent récapitulent l'ensemble des étapes suivies et des calculs faits pour aboutir à la distance K-S :

Quantile	Nombre d'observations	Nombre cumulé des observations	Fonction de répartition empirique	Quantile centré réduit	Fonction de répartition théorique normale	La distance D (cas normale)
0	0	0	0	-1,11	0,1335	-0,1335
500	15	15	0,227272727	-0,79	0,21476	0,01251273
1000	17	32	0,484848485	-0,47	0,31918	0,16566848
1500	6	38	0,575757576	-0,15	0,44038	0,13537758

2000	7	45	0,681818182	0,17	0,56749	0,11432818
2500	1	46	0,696969697	0,49	0,68793	0,0090397
3000	6	52	0,787878788	0,80	0,78814	-0,00026121
3500	4	56	0,848484848	1,12	0,86864	-0,02015515
4000	3	59	0,893939394	1,44	0,92507	-0,03113061
4500	3	62	0,939393939	1,76	0,9608	-0,02140606
5000	1	63	0,954545455	2,08	0,98124	-0,02669455
5500	1	64	0,96969697	2,40	0,9918	-0,02210303
6000	1	65	0,984848485	2,72	0,99674	-0,01189152
6500	1	66	1	3,04	0,99882	0,00118

Quantile	Nombre d'observations	Nombre des observations cumulé	Fonction de répartition empirique	Fonction de répartition théorique Gamma	La distance D (cas Gamma)
0	0	0	0	0	0
500	15	15	0,227272727	0,2005618	-0,026710927
1000	17	32	0,484848485	0,3953551	-0,089493385
1500	6	38	0,575757576	0,5510172	-0,024740376
2000	7	45	0,681818182	0,6700236	-0,011794582
2500	1	46	0,696969697	0,7591166	0,062146903
3000	6	52	0,787878788	0,8249974	0,037118612
3500	4	56	0,848484848	0,8733187	0,024833852
4000	3	59	0,893939394	0,9085564	0,014617006
4500	3	62	0,939393939	0,9341424	-0,005251539
5000	1	63	0,954545455	0,9526582	-0,001887255
5500	1	64	0,96969697	0,9660217	-0,00367527
6000	1	65	0,984848485	0,9756456	-0,009202885
6500	1	66	1	0,9825638	-0,0174362

Annexe 6: Les matrices variance-covariance des paramètres des modèles GLM

La matrice variance-covariance des paramètres du modèle GLM s'obtient en ajoutant à la procédure GENMOD l'option Covb. Prenons l'exemple de la modélisation poisson :

```
proc genmod data=sasuser.regression_1;
model col1=A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11
/ dist= poisson link=log covb corrb;
ods output ParameterEstimates=parametre covb=var_cov_para_poisson;
run;
```

	MU	A0	A2	A3	A4	A5		B6	B7	B8	B9	B10
MU	0,251086	-0,00109	-0,001	-0,001	-0,001	-0,001		-0,25007	-0,25007	-0,25006	-0,25006	-0,25005
A0	-0,00109	0,001086	0,001003	0,001003	0,001002	0,001002		6,92E-05	6,64E-05	6,3E-05	5,63E-05	4,56E-05
A2	-0,001	0,001003	0,001071	0,001003	0,001002	0,001002		-1,4E-05	-1,7E-05	-2E-05	-2,7E-05	-3,7E-05
A3	-0,001	0,001003	0,001003	0,001097	0,001002	0,001002		-1,4E-05	-1,7E-05	-2E-05	-2,7E-05	-3,9E-14
A4	-0,001	0,001002	0,001002	0,001002	0,001086	0,001002		-1,4E-05	-1,7E-05	-2,1E-05	-3,9E-14	-3,9E-14
A5	-0,001	0,001002	0,001002	0,001002	0,001002	0,001112		-1,5E-05	-1,8E-05	-3,9E-14	-3,9E-14	-3,9E-14
B6	-0,25007	6,92E-05	-1,4E-05	-1,4E-05	-1,4E-05	-1,5E-05		0,250377	0,250066	0,250063	0,250056	0,250046
B7	-0,25007	6,64E-05	-1,7E-05	-1,7E-05	-1,7E-05	-1,8E-05		0,250066	0,250506	0,250063	0,250056	0,250046
B8	-0,25006	6,3E-05	-2E-05	-2E-05	-2,1E-05	-3,9E-14		0,250063	0,250063	0,250998	0,250056	0,250046
B9	-0,25006	5,63E-05	-2,7E-05	-2,7E-05	-3,9E-14	-3,9E-14		0,250056	0,250056	0,250056	0,251406	0,250046
B10	-0,25005	4,56E-05	-3,7E-05	-3,9E-14	-3,9E-14	-3,9E-14		0,250046	0,250046	0,250046	0,250046	0,291712

Le même traitement est refait pour la modélisation Gamma. Le marice se présente comme suit :

	MU	A0	A2	A3	A4	A5		B6	B7	B8	B9	B10
MU	0,364976	-0,20052	-0,1849	-0,18097	-0,18036	-0,18012		-0,18053	-0,17873	-0,17839	-0,17367	-0,17086
A0	-0,20052	0,200524	0,184901	0,180971	0,180358	0,180121		0,016077	0,01428	0,013941	0,009221	0,006404
A2	-0,1849	0,184901	0,203993	0,1818	0,18034	0,179969		0,001495	-0,00062	-0,01036	-0,01084	-0,00783
A3	-0,18097	0,180971	0,1818	0,199149	0,180209	0,179717		-0,00341	-0,00208	-0,00442	-0,00076	-0,00034
A4	-0,18036	0,180358	0,18034	0,180209	0,200572	0,179651		-0,00415	-0,00364	-0,0005	1,24E-05	7,24E-06
A5	-0,18012	0,180121	0,179969	0,179717	0,179651	0,203059		-0,00323	-0,00573	0,00019	9,32E-05	6,24E-05
B6	-0,18053	0,016077	0,001495	-0,00341	-0,00415	-0,00323		0,207424	0,17828	0,177771	0,173078	0,17043
B7	-0,17873	0,01428	-0,00062	-0,00208	-0,00364	-0,00573		0,17828	0,210551	0,177215	0,17321	0,17056
B8	-0,17839	0,013941	-0,01036	-0,00442	-0,0005	0,00019		0,177771	0,177215	0,224169	0,178605	0,174416
B9	-0,17367	0,009221	-0,01084	-0,00076	1,24E-05	9,32E-05		0,173078	0,17321	0,178605	0,230868	0,172675
B10	-0,17086	0,006404	-0,00783	-0,00034	7,24E-06	6,24E-05		0,17043	0,17056	0,174416	0,172675	0,252512

Annexe 7 : Le code de la macro SAS du Bootstrap appliqué au modèle Gamma

```
proc iml;
mat=j(1,1000,.);
create table_resume from mat;
append from mat;
close table_resume;
quit;
%macro boot(b,table);
/* Bootstraper une fois les résidus*/
proc surveyselect data= vect_residus
    out= out
    method=URS
    samprate=1
    outhits noprint
    rep=1;
run;
/* Construction du vecteur des increments*/
proc iml;
use out;
read all var _all_ into rsd_boot;
use vect_gamma_increm;
read all var _all_ into pred;
mat=j(66,1,0);
do i=1 to 66;
mat[i,1]=rsd_boot[i,1]*1590+pred[i,1];
if mat[i,1]<0 then mat[i,1]=0;
end;
print mat;
create vect_ from mat;
append from mat;
close vect_;
quit;
/*Creation du tableau regression*/
data regression_1;
merge vect_ facteurs1;
run;
proc print data=regression_1;
run;
/*estimation des paramètres de régression*/
proc genmod data=regression_1;
model col1=A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11
/ dist= gamma link=log;
ods output ParameterEstimates=parametre;
run;
proc print data=parametre;
run;
data estim_coef_reg(keep=Estimate);
set parametre;
run;
/*Construction du tableau complet gamma*/
proc iml;
use estim_coef_reg; read all var _all_ into mat;
use facteurs; read all var _all_ into x;
beta=j(23,1,0);
do i=1 to 23;
beta[i,1]=mat[i,1];
end;
print x;
print beta;
X= x * beta;
y=exp(x);
print y;
Create vect_gamma from y;
append from y;
close vect_gamma;
quit;
proc iml;
use vect_gamma;
read all var _all_ into mat1;
mat=j(11,11,.);
k=1;
do i=1 to 11;
```

```

do j=1 to 11;
mat[i,j]=mat1[k,1];
k=k+1;
end;
end;
print mat;
create tabl_compl_gamma from mat;
append from mat;
close tabl_compl_gamma;
quit;
proc print data=tabl_compl_gamma;
run;
/*calcul de la reserve globale pour le tableau de gamma*/
data tabl_compl_gamma_cumul( drop=i);
set tabl_compl_gamma;
array cumul [*] _numeric_;
do i=2 to 11;
cumul(i)=cumul(i-1) + cumul(i);
end;
run;
proc iml;
use tabl_compl_gamma_cumul;
read all var _all_ into mat;
mat1=j(11,1,0);
k=1;
do i=1 to 11;
do j=1 to 11;
if i+j=12 then mat1[k,1]=mat[i,j];
if i+j=12 then k=k+1;
end;
end;
create vect from mat1;
append from mat1;
close vect;
quit;
data vect;
set vect;
rename col1=diag;
run;
data tabl_compl_gamma_cumul;
merge tabl_compl_gamma_cumul vect;
run;
data tabl_compl_gamma_cumul;
set tabl_compl_gamma_cumul;
reserv=col11-diag;
run;
proc sql;
create table reserv_global_gamma as select sum(reserv) as total_provision
from tabl_compl_gamma_cumul ;
quit;
proc iml;
use reserv_global_gamma;
read all var _all_ into mat1;
use &table;
read all var _all_ into mat;
mat[1,&b]=mat1[1,1];
create table_resume_1 from mat;
append from mat;
close table_resume_1;
quit;
proc iml;
use &table; read all var _all_ into mat;
use table_resume_1; read all var _all_ into mat1;
do i=1 to 1000;
mat[1,i]=mat1[1,i];
end;
create table__ from mat;
append from mat;
close table__;
quit;
proc iml;
use table__;
read all var _all_ into mat;
create &table from mat;

```

```
append from mat;  
close &table;  
quit;  
%MEND;
```

