CEA | CENTRE D'ÉTUDES ACTUARIELLES

LOIS BEST ESTIMATE POUR LE MODELE A ETATS D'UN PORTEFEUILLE TNS EN INCAPACITE/INVALIDITE

Audrey INGENITO BERTAL

Directeur de Mémoire
Philippe SEIGNEUR

CEA | CENTRE D'ÉTUDES ACTUARIELLES

INSTITUT DES ACTUAIRES

Mémoire présenté devant le Centre d'Etudes Actuarielles pour la validation du cursus à la Formation d'Actuaire du Centre d'Etudes Actuarielles et l'admission à l'Institut des Actuaires le 29 mars 2016

Par : Audrey INGENITO BERTAL

Titre : Lois Best Estimate pour le modèle à états d'un portefeuille de TNS en incapacité/invalidité

Confidentialité : ☐ NON ☒ OUI (Durée : ☐ 1an ☒ 2 ans)
Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

Membre présent du jury de l'Institut des Actuaires :

_____

_____

_____

Membres présents du jury du Centre d'Etudes Actuarielles :

_____

_____

_____

_____

_____

_____

_____

_____

Secrétariat :

Bibliothèque :

Entreprise :

Nom : GENERALI

Signature :

Directeur de mémoire en entreprise :

Nom : M. Philippe SEIGNEUR

Signature :

Invité :

Nom : _____

Signature :

Autorisation de publication et de mise en ligne sur un site de diffusion de documents actuariels
(après expiration de l'éventuel délai de confidentialité)

Signature du responsable entreprise

Signature(s) du candidat(s)

SOMMAIRE

Resume 6

Abstract 7

Remerciements 9

Introduction 11

I. PRESENTATION DU PORTEFEUILLE 13

A. PRESENTATION DU MARCHE DES TNS 13

B. DIFFERENCES DE REGIME ENTRE SALARIES ET TNS 14
1. L'incapacité des salariés vs l'incapacité des TNS 14
2. L'invalidité des salariés vs l'invalidité des TNS 14

C. PRESENTATION DU PORTEFEUILLE ETUDIE 15
1. PGM/PPL 15
   a. Le risque incapacité 16
   b. Le risque invalidité 16
2. ATOLL 17
   a. Le risque incapacité 17
   b. Le risque invalidité 19

D. STATISTIQUES DESCRIPTIVES DU PORTEFEUILLE 19
1. Statistiques sur le portefeuille PPL/PGM 19
2. Statistiques sur le portefeuille ATOLL (toute population) 20
3. Comparaison du risque incapacité/invalidité entre notre portefeuille de TNS et notre portefeuille de salariés 20

II. MODELE A ETATS ET QDD CHEZ GENERALI 23

A. PRESENTATION DU MODELE A ETATS 23

B. LA QUALITE DES DONNEES CHEZ GENERALI 26

III. CONSTRUCTION DES TABLES DE MAINTIEN EN INCAPACITE, DE PASSAGE EN INVALIDITE ET DE MAINTIEN EN INVALIDITE 31

A. LES DONNEES 31
1. Les bases de données de l'ancien système 32
   a. Les fichiers sources 32
   b. Le traitement des données 32
2. Les bases de données du système de gestion actuel 35
   a. Les tables sources 36
   b. Le traitement des données 37
3. La base consolidée 40
   a. Le retraitement des données 40
   b. L'agrégation des lignes 42
   c. Le contrôle sur la qualité des données retraitées 43
   d. Le traitement des rechutes 43
   e. La franchise incapacité 44
   f. L'incapacité dépassant les 36 mois 45
   g. Le récapitulatif des retraitements 45
4. Les statistiques sur les données finales 47
   a. Les statistiques liées à l'incapacité 47
   b. Les statistiques liées à l'invalidité 51

B. LA CONSTRUCTION DE LA TABLE DE MAINTIEN EN INCAPACITE 54
1. Les données 54
   a. Les troncatures 55

b. Les censures 55
c. La table finale des données 55
2. L'estimateur de Kaplan Meier 57
a. Notations 57
b. Obtention de l'estimateur 58
c. Application de l'estimateur 60
3. L'obtention des taux bruts 61
a. Résultat 61
b. Intervalles de confiance de l'estimateur 61
4. L'obtention des taux lissés 62
a. Méthode du lissage de Whittaker Henderson 62
b. Validation du lissage 64
c. Résultat 65

C. LA CONSTRUCTION DE LA TABLE DE PASSAGE EN INVALIDITE 67
1. Les données 67
a. Les variables utiles 68
b. Le tableau récapitulatif 68
c. Le volume des observations 68
2. L'obtention des taux bruts 69
3. L'obtention des taux lissés 70
a. Choix des paramètres pour le lissage de Whittaker Henderson 71
b. Validation du lissage 71
c. Résultat 71

D. LA CONSTRUCTION DE LA TABLE DE MAINTIEN EN INVALIDITE 75

E. IMPACTS SUR LE CALCUL DES PROVISIONS 77
1. Le calcul des provisions techniques 77
a. Notations 77
b. Le calcul des PSAP 77
c. Le calcul de l'invalidité en attente 78
d. Le calcul des PMR 79
e. Le coefficient d'abattement pour l'incapacité 79
2. La comparaison des montants de provisions 80
a. L'impact sur les PSAP 80
b. L'impact sur les PSAPR 81
c. L'impact sur les PMR 82
d. L'impact réel sur notre portefeuille 83

IV. CONSTRUCTION DES LOIS D'ENTREE EN INCAPACITE 87

A. LES DONNEES 87
1. La présentation des fichiers sources 87
2. Les traitements des données 88
a. Préparation des bases 88
b. Traitement de la base des sinistres 88
c. Construction des tables de bases 88
d. Retraitements des données 89
3. Les statistiques sur le portefeuille obtenu 89
a. Nombre de contrats en cours 89
b. Nombre de nouveaux sinistres 91

B. LA METHODE 93
1. Les taux bruts 93
2. Le lissage 96
a. Méthode utilisée 96
b. Courbe lissée 97
c. Choix du lissage 98
d. Reconstitution d'une courbe lissée par âge 99

C. LES RESULTATS **99**

V. CONSTRUCTION DES LOIS DE CHUTES ***101***

A. LES DONNEES **102**
1. La présentation des bases de données 102
2. Le traitement des données 103
a. Nettoyage des bases 103
b. Traitement des bases 103
c. Consolidation des bases 103
d. Retraitements des données 104
3. Statistiques sur le portefeuille obtenu 104

B. LA METHODE **105**
1. Les taux bruts 105
2. Le lissage 107
a. Choix du lissage pour PGM/PPL 107
b. Choix du lissage pour ATOLL 108

C. LES RESULTATS **109**
1. Résultats PGM/PPL 109
2. Résultats ATOLL TNS 110

Conclusion ***113***

Annexes ***115***

Bibliographie ***125***

RESUME

Dans le cadre de Solvabilité 2, Generali a choisi d'utiliser un modèle interne pour calculer son risque en prévoyance individuelle. Ainsi, notre étude porte sur la construction des lois incapacité/invalidité des TNS permettant l'alimentation des hypothèses du modèle à états développé. Les lois construites permettent de faire évoluer l'état des assurés à chaque pas de projection, de calculer les provisions en Best Estimate et les cash-flows futurs et de calibrer les chocs à appliquer pour le calcul du capital économique.

Après avoir contrôlé et retraité les données, nous avons utilisé l'estimateur de Kaplan Meier et un lissage de Whittaker Henderson pour obtenir une table de maintien en incapacité par âge et par ancienneté. Nous avons également construit une table de passage d'incapacité en invalidité. Par manque de données, nous n'avons pas pu établir une loi de maintien en invalidité. Pour évaluer les provisions du portefeuille en Best Estimate, nous utilisons nos tables d'expérience et la table de maintien en invalidité que le BCAC est en train de finaliser.

Nous construisons également une loi d'incidence en incapacité pour modéliser la fréquence d'entrée en arrêt de travail par âge et par produit. La courbe brute est lissée par une régression polynomiale.

Enfin, nous évaluons un taux de chute par ancienneté et par produit à l'aide de la méthode Chain-Ladder. Nous lissons la loi brute par une régression polynomiale.

L'ensemble de ces lois permettent, d'alimenter le modèle interne avec des données reflétant le comportement du portefeuille TNS de Generali.

***MOTS CLES** : incapacité, invalidité, TNS, tables d'expériences, hypothèses modèle interne, incidence, maintien, passage, chutes, provisions*

ABSTRACT

Generali chose to use an internal model to calculate individual protection business risk for the Solvency 2 directive. This study concerns the setup of temporary / permanent disability, for a self-employed worker portfolio, to give actuarial assumptions to the state model. The laws have been put in place to either to define the policyholder number, depending on their state for each predicted period, or for the calculation of Best Estimate reserves and the future cash flows, or to calibrate shocks which should be applied to the model for SCR calculating.

After data control and data processing, the Kaplan Meier estimator and the Whittaker Henderson adjustment method are used to establish the recovery of temporary disability table, for each age and duration. We have also created a table for the migration from temporary to permanent disability. We can't obtain a reliable table for the recovery of permanent disability because the data base is not sufficient. We are using our experience tables and the duration of permanent disability table updated at the moment with the BCAC to calculate the Best Estimate reserves.

We have also created an incidence of temporary disability table according to the product and the age. A polynomial regression is used for the curve smoothing.

Finally we have set up a table for lapses with the Chain-Ladder method according to the seniority and the product. The table results are adjusted by a polynomial regression.

The tables resulting from this thesis are used now to model the temporary / permanent disability risk of the Generali self-employed workers portfolio.

***Keywords** : Temporary and permanent disability, self-employed workers, experience tables, assumptions to the internal model, incidence, recovery, migration, lapses, reserves*

REMERCIEMENTS

Je tiens à remercier Monsieur Christophe IZART, Directeur du Plan et Contrôle de Gestion et Monsieur Pierre CORREGE Directeur Provisionnement et Valeur, pour leurs remarques constructives et le temps qu'ils m'ont consacré.

Je remercie Monsieur Philippe SEIGNEUR, Manager du Service Provisionnement Prévoyance et Santé, pour m'avoir proposé ce sujet d'étude. Je le remercie pour ses conseils et son suivi. Ses recommandations m'ont permis d'avancer et de mener à bien cette étude.

Mes remerciements vont également à Monsieur Babacar M'BAY et Monsieur Cédric JACQUET, pour avoir répondu à mes différentes sollicitations.

Enfin, je tiens également à remercier mes collègues du Service Provisionnement ainsi que tous mes proches pour leurs soutiens et leurs encouragements tout au long de ma formation au CEA.

INTRODUCTION

Aujourd'hui avec Solvabilité 1, les hypothèses fixes utilisées pour un calcul unique de provisions comportent une marge de prudence. Cette vision, qui ne laisse pas de place à l'incertitude n'appréhende pas le risque de manière optimale. Demain avec Solvabilité 2, la provision sera la somme d'une provision calculée en Best Estimate et d'une marge de risque explicite. La meilleure estimation des provisions correspond à la moyenne pondérée par leur probabilité des flux de trésorerie futurs calculés pour les différents scénarios sur la base d'informations actualisées, crédibles et d'hypothèses réalistes, notamment des lois d'expérience spécifiques au portefeuille concerné.

Generali a ainsi mis en place sous Prophet un modèle interne « à états » utilisant des paramètres spécifiques pour les produits dédiés aux Travailleurs Non-Salariés (TNS), portefeuille étudié dans ce mémoire. L'assureur souhaite également avoir des statistiques précises et actualisées sur le risque arrêt de travail pour mieux modéliser son risque. Ainsi, les lois d'expérience éviteront de surévaluer ou sous-estimer les provisions et de prendre en compte la sinistralité réelle du portefeuille étudié pour les projections. En 2008, des tables d'expériences pour le maintien en incapacité et le passage en invalidité ont été construites sur ce portefeuille par un cabinet d'actuaires en mutualisant les données TNS de plusieurs sociétés. Nous avons décidé d'actualiser ces tables en prenant en compte les nouvelles données disponibles.

Nous exposerons, tout d'abord, le portefeuille TNS étudié avec une présentation des produits avec quelques statistiques descriptives ainsi qu'un aperçu de ce marché spécifique.

Ensuite, nous évoquerons succinctement le modèle interne et la qualité des données mis en place sur la prévoyance individuelle chez Generali.

Puis, la troisième partie sera consacrée à l'établissement des tables d'expérience : loi de maintien en incapacité, loi de passage en invalidité et loi de maintien en invalidité (si assez de données). Nous retraiterons les données disponibles pour nettoyer les données et avoir une base propre et cohérente pour notre étude permettant d'avoir une qualité des données suffisante comme exigée par Solvabilité 2. Nous construirons alors des lois brutes puis des lois lissées et nous vérifierons les impacts sur le montant des provisions.

La quatrième partie décrira la construction du taux d'incidence en incapacité sur le portefeuille.

Et enfin la dernière partie présentera l'établissement des lois de chutes.

L'ensemble de ces lois permettra de calculer les provisions en Best Estimate, de déterminer les différents chocs à appliquer au modèle et d'avoir les probabilités de passage d'un état à l'autre avec leurs flux correspondants pour le modèle prévoyance TNS pour l'application de la directive Solvabilité 2 au 1er janvier 2016.

I. PRESENTATION DU PORTEFEUILLE

Dans un premier temps, il convient de présenter le marché spécifique des TNS, les produits étudiés ainsi que certaines statistiques du portefeuille.

A. PRESENTATION DU MARCHE DES TNS

La totalité de notre étude porte sur des contrats souscrits par des TNS (Travailleurs Non-Salariés) qui, en souscrivant des contrats dits « Madelin » (loi n° 94-126 du 11 février 1994), bénéficient d'avantages fiscaux en déduisant de leur revenu imposable les cotisations versées chaque année pour un contrat de retraite complémentaire, de garanties de prévoyance ou de santé.

Les travailleurs indépendants ne disposent pas de la même couverture sociale que les travailleurs salariés. Ils ne cotisent pas aux mêmes organismes, que ce soit pour la santé, la retraite ou la prévoyance. Ces assurés doivent, en majorité, être affiliés au régime obligatoire RSI (Régime Social des Indépendants) et ils souscrivent des garanties complémentaires, car, en générale, les prestations de base sont insuffisantes par rapport aux revenus perçus. Cette sécurité sociale des indépendants a été créée en 2006 par la fusion des caisses de protection sociale des chefs d'entreprises. Elle effectue une mission de service public en gérant la protection sociale obligatoire de plus de 5,6 millions de personnes protégées (2/3 d'assurés et 1/3 d'ayants droit), ce qui le place comme deuxième régime de protection sociale en France par le nombre de personnes assurées. Les cotisants du RSI sont pour plus de 2/3 des hommes alors que sur l'ensemble de la population active leur proportion est de 1/2. Le RSI encaisse 22,9 Mds € de produits (dont 11Mds € de cotisations et 7,5 Mds € de contributions sociales) et versent 22 Mds € de charges (dont 17 Mds de prestations au total et 664 millions € au titre de l'incapacité/invalidité).

Concernant la prévoyance Madelin à fin 2014 en France, les cotisations s'élèvent à 2,1 Mds € (dont 47% concerne les garanties incapacité/invalidité) soit une progression modérée de 3% par rapport à 2013, pour 1,9 millions de contrats en cours. Ces contrats sont détenus par 1,8 millions d'assurés, qui, rapportés au nombre d'indépendants actifs couverts par le RSI, placent le taux d'équipement à 79%. Ce taux d'équipement est cependant très variable selon le type de garanties souscrites. En effet, si 4 indépendants sur 10 sont couverts en frais de soins ou en incapacité-invalidité, seulement 1 sur 4 l'est en décès.
La prime annuelle totale moyenne s'établit environ à 1 100€ dont 990€ pour la garantie incapacité/invalidité. La charge des prestations de 1,2 Mds € enregistre une faible baisse de 2% en 2014.

Sur le marché français, le portefeuille TNS de Generali représente 6,4% des contrats en cours (soit au 5^{e} rang) et 10,7% des cotisations payées (soit en 3^{e} position).

B. DIFFERENCES DE REGIME ENTRE SALARIES ET TNS

1. L'incapacité des salariés vs l'incapacité des TNS

Lorsque l'assuré salarié se trouve en arrêt de travail, la sécurité sociale lui octroie des indemnités journalières (IJ). Elles sont égales à 50 % du salaire journalier de base (dans la limite du plafond de la sécurité sociale (PASS) soit 43€ pour 2015). Sauf exception, le versement des IJ est soumis à un délai de franchise de 3 jours. La durée du versement est limitée à 3 ans. Au-delà, si l'assuré n'a pas repris le travail, il passe obligatoirement en état d'invalidité.
Certaines conventions collectives permettent que le salarié (sous certaines conditions), bénéficie, en cas d'incapacité, d'une indemnité complémentaire aux IJ de la sécurité sociale pour maintenir le montant intégrale ou en partie du salaire pendant l'arrêt de travail. Cette indemnité complémentaire est versée par l'employeur qui a l'obligation légale, dès le 1er jour d'arrêt (en cas d'accident du travail ou de maladie professionnelle) ou passé un délai de franchise de 7 jours sinon, de maintenir le salaire de son employé à un certain niveau et pendant une durée qui dépend de l'ancienneté de l'employé dans l'entreprise. L'employeur déduit donc du maintien de salaire, les IJ de la sécurité sociale perçues par le salarié afin que sa rémunération ne soit pas supérieure à celle qu'il touche lorsqu'il travaille.

En cas d'arrêt de travail d'un TNS, le RSI lui verse, sous certaines conditions, une IJ à partir du 4e jour en cas d'hospitalisation et du 8e jour en cas de maladie ou accident. Son montant est égal à 50% du revenu professionnel moyen des 3 dernières années dans la limite du PASS (soit un maximum de 52 € par jour pour l'année 2015). La durée du versement est limitée à 360 jours d'indemnisation sur une période de 3 ans sauf pour les affections de longues durées ou il est possible de bénéficier de 3 années de versement au maximum sous réserve que l'arrêt de travail soit médicalement justifié.

2. L'invalidité des salariés vs l'invalidité des TNS

Concernant la rente d'invalidité, elle débute :

- soit à l'expiration de la durée maximale de perception des IJ, soit 3 ans
- soit à la date de stabilisation de l'état de santé

La rente d'invalidité versée par la sécurité sociale est calculée sur la base d'une rémunération moyenne (dans la limite du PASS). Le montant de la rente versée est égal à cette base multipliée par le pourcentage dépendant de la catégorie d'invalidité (soit une rente mensuelle comprise entre 282€ au minimum et 1 585€ + 1 103€ de majoration pour une tierce personne au maximum). La prestation cesse d'être versée au moment du départ légal à la retraite.

Le versement des rentes invalidité du RSI diffère suivant les professions. Pour un artisan par exemple, sous certaines conditions, le RSI lui verse une prestation au bout du 91ème jour d'arrêt. La rente annuelle est égale à 50 % du revenu annuel moyen pour une invalidité totale et définitive à toute activité. Cette pension peut être versée jusqu'à l'âge de 60 ans au maximum.

En règle générale, les TNS ont des régimes sociaux qui octroient des niveaux de prestations bien inférieurs à ceux du régime de la sécurité sociale. D'autre part, il n'y a pas de « maintien de salaire » comme pour les salariés. Les TNS ont tout intérêt, s'ils désirent être mieux couverts pour protéger leur famille des conséquences d'une perte soudaine de revenus, à souscrire une assurance prévoyance individuelle. Ils bénéficient ainsi d'une couverture complémentaire en cas d'arrêt de travail et ils choisissent le montant, la durée et le début de versement du complément de salaire en cas de sinistre. L'utilisation des tables de provisionnement réglementaires du BCAC (construites sur une population de salariés) n'est à priori pas adéquate car elle ne reflète pas le comportement de cette population spécifique.

C. PRESENTATION DU PORTEFEUILLE ETUDIE

Le présent mémoire étudie le comportement du portefeuille de prévoyance incapacité/invalidité des TNS de Generali. Ce portefeuille étant principalement composé de deux produits, nous présentons les caractéristiques de chacun puis nous comparerons le risque incapacité/invalidité de notre portefeuille TNS et de notre portefeuille salariés.

1. PGM/PPL

Les produits PGM (Plan Gérant Majoritaire) et PPL (Plan Profession Libérale) ont été lancés par Generali en 1989 avec, depuis, plusieurs versions de produits.
Ce produit a pour objet de garantir en fonction du choix effectué par l'assuré :

- Un capital en cas de décès ou de perte totale et irréversible d'autonomie
- Une rente de conjoint et/ou d'éducation en cas de décès de l'assuré
- Des indemnités en cas d'incapacité de travail ou d'invalidité
- Le remboursement des frais médicaux
- Une exonération des cotisations en cas d'incapacité ou d'invalidité

Les garanties obligatoires sont le capital décès et/ou la rente de conjoint.
Notre étude s'intéressera exclusivement aux garanties annexes concernant l'incapacité et l'invalidité.

a. Le risque incapacité

En cas d'arrêt de travail, pour compléter le versement du RSI et diminuer cette perte de revenu, Generali verse à l'assuré une IJ dont le montant est fixé à l'adhésion en fonction du PASS en vigueur le jour de l'arrêt de travail.

L'indemnité journalière est payable à l'expiration d'un délai de franchise de :

- 30 jours d'arrêt de travail continu, en cas de maladie
- 3 jours d'arrêt de travail continu en cas d'accident ou d'hospitalisation

Aucune indemnité n'est versée en cas d'incapacité temporaire partielle.

Si un assuré, après avoir repris son travail pendant une durée inférieure à 60 jours, est victime d'un nouvel arrêt de travail, il n'y aura pas application du délai de franchise et les prestations seront versées immédiatement (nous parlons alors de rechute). Par contre, en cas de nouvel arrêt après une reprise de travail d'une durée égale ou supérieure à 60 jours, le délai de franchise sera à nouveau appliqué.

L'indemnité est payable pendant toute la durée de l'incapacité. Elle cesse d'être due :

- dès la reprise, même partielle, de travail
- en cas de décès de l'assuré
- et au plus tard, au 65^{e} anniversaire de l'assuré

b. Le risque invalidité

En cas d'invalidité permanente partielle ou totale d'un assuré, Generali lui règle une rente dont le montant annuel est déterminé à l'adhésion en fonction du PASS et suivant le taux d'invalidité déterminé par expertise entre 51% (Invalidité Permanente Partielle) et 85% (Invalidité Permanente Totale).

La rente cesse d'être due :

- si le taux d'invalidité devient inférieur à un tiers
- en cas de liquidation de la pension de vieillesse
- en cas de décès
- au plus tard au 65^{e} anniversaire de l'assuré (l'âge étant spécifié, cela limite le risque lié à une révision de l'âge légal par l'Etat de départ en retraite)

2. ATOLL

Le produit ATOLL est commercialisé par Generali depuis 2004 avec depuis plusieurs générations de produits. Ce contrat peut être souscrit par des salariés ou Travailleurs Non-Salariés. Ce produit possède un choix de garanties plus large que PGM/PPL.
Les différentes garanties proposées à l'assuré sont :

- Un capital en cas de décès ou de perte totale et irréversible d'autonomie
- Une rente viagère de conjoint et/ou une rente éducation
- Un capital et/ou une rente en cas d'invalidité
- Une IJ en cas d'incapacité temporaire totale ou partielle
- Une IJ cas d'incapacité temporaire totale après hospitalisation
- Des frais professionnels
- Un relais régime professionnel
- Une exonération des cotisations en cas d'incapacité/invalidité

Notre étude s'intéressera exclusivement aux garanties indemnités journalières en incapacité et rentes invalidité des assurés TNS.
La population TNS pouvant souscrire ce produit est constituée des professions médicales/paramédicales, des professions libérales, des commerçants, des artisans, des professions expert/conseils et des professions agricoles.

a. Le risque incapacité

La plupart des garanties incapacité d'ATOLL peuvent se cumuler pour un même sinistre survenu sur des durées différentes et avec des montants différents. C'est-à-dire que pour la survenance d'un seul sinistre, plusieurs garanties sont « activées » et l'assuré percevra plusieurs prestations en même temps pour couvrir par exemples sa perte de revenu mais également les frais généraux de sa société.

i. Incapacité Temporaire Totale : ITT

En cas d'ITT de l'assuré, les IJ sont versées après expiration du délai de franchise, défini au certificat d'adhésion. Elles correspondent à un revenu de remplacement.

Dans la dernière version d'ATOLL, la garantie ITT se présente sous forme de 3 paliers d'IJ dont le montant peut être différent :

- le 1er palier couvre, selon la franchise appliquée, jusqu'au 90ème jour
- le 2ème palier couvre à partir du 91ème jour et jusqu'au 365ème jour
- le dernier palier couvre du 366ème jour et jusqu'au 1095ème jour

L'ancienne version du produit est « classique » c'est à dire sans palier avec un montant d'IJ constant sur les 3 ans et avec un large choix de franchises possibles.

Les indemnités sont versées tant que dure l'incapacité et cesse :

- à la date de liquidation de ses droits à la retraite par son régime général
- à la consolidation d'un éventuel état d'invalidité
- au bout des 3 années de prestations versées
- au plus tard, à la date d'échéance annuelle de l'adhésion qui suit soit le 65ème anniversaire (pour les anciennes versions du produit) soit le 70ème anniversaire (pour la version la plus récente d'ATOLL) de l'assuré

ii. Incapacité Temporaire Partielle : ITP

Si, à l'issue de l'ITT ayant donné lieu à indemnisation, l'assuré peut reprendre partiellement son activité professionnelle, l'indemnisation pourra être poursuivie jusqu'à la reprise complète de son activité et pendant une durée maximale de 6 mois. L'indemnité est égale à 50 % de l'indemnité versée au titre de l'ITT.

iii. Incapacité Temporaire Totale après hospitalisation

Cette garantie ne peut être souscrite qu'en complément à la garantie indemnités journalières en cas d'ITT.
En cas d'ITT de l'assuré, faisant suite à une hospitalisation, le montant de l'IJ souscrit est réglé à l'expiration du délai de franchise hospitalisation prévu au certificat d'adhésion, et pendant la durée de l'hospitalisation et de l'arrêt de travail consécutif à celle-ci, sans pouvoir excéder 365 jours.

iv. Frais Professionnels

Cette garantie permet de couvrir les frais généraux de la société de l'assuré ou de son activité professionnelle (salaires des employés, loyers, etc…) en cas d'ITT ou d'ITP. Les IJ définies au certificat d'adhésion sont versées après l'expiration du délai de franchise selon la durée de garantie choisie : soit jusqu'au 365ème jour soit jusqu'au 1095ème jour.
Cette garantie peut se cumuler avec la garantie ITT ou ITP.

v. Relais régime professionnel

En cas d'ITT, l'assuré reçoit, après expiration du délai de franchise, les IJ définies au certificat d'adhésion. Les prestations sont versées tant que dure l'ITT et au plus tard jusqu'au 90ème jour d'arrêt de travail.

b. Le risque invalidité

i. Capital invalidité

L'assuré atteint d'invalidité permanente fonctionnelle peut demander, dans les 2 mois à partir de la date de consolidation, le règlement du capital invalidité dont le montant est défini à l'adhésion.

ii. Rente invalidité

En cas d'invalidité permanente d'un taux supérieur ou égal à 66 %, l'assuré perçoit une rente mensuelle indiquée au certificat d'adhésion et cesse :

- au moment où il bénéficie d'une pension de retraite versée par le régime général
- au plus tard à la date d'échéance annuelle de l'adhésion qui suit ses 65 ans

Lorsque le taux d'invalidité est supérieur à 33 % mais inférieur à 66 %, la rente est versée dans les mêmes conditions en lui appliquant le rapport du taux constaté à 66 %.

D. STATISTIQUES DESCRIPTIVES DU PORTEFEUILLE

Il est important de relever qu'au sein même de notre portefeuille d'assurés TNS, suivant le produit, la population est différente. En effet, en étudiant les professions des contrats en cours à fin 2013, nous observons que PGM/PPL concentre surtout des chefs d'entreprises, des professions libérales (consultant, expert-comptable, architecte …) et des artisans alors qu'ATOLL est composé surtout de professions libérales médicales (infirmier, kinésithérapeute, dentiste…), des artisans et des commerçants. Cette différence de catégorie socio-professionnelle peut engendrer des comportements et des sinistralités différentes notamment sur l'incidence par exemple.

1. Statistiques sur le portefeuille PPL/PGM

Au 31/12/2013, 26 859 contrats en cours pour un chiffre d'affaires de 75 millions d'euros (dont 51 millions pour les garanties incapacité/invalidité) et une prime moyenne annuelle de 2 620€.
2 105 affaires nouvelles ont été souscrites pour un montant d'APE (Annual Premium Equivalent) de 4,5 millions d'euros.
Les assurés ont en moyenne 51 ans, soit 6 ans de plus que l'âge moyen des cotisants du RSI; ils sont également plus âgés que les salariés qui ont une moyenne d'âge de 40 ans.
Le coût moyen d'un sinistre incapacité (cumul des règlements et des provisions) est de 116 876€ pour une IJ moyenne de 154€ et une durée moyenne de 8,4 mois. L'incapable a en moyenne 53 ans.
La charge des prestations est de 82 millions d'euros (dont 77 millions pour l'incapacité/invalidité).

2. Statistiques sur le portefeuille ATOLL (toute population)

Au 31/12/2013, 56 857 contrats en cours pour un chiffre d'affaires de 66 millions d'euros (dont 49 millions pour les garanties incapacité/invalidité) et une prime moyenne annuelle de 1 225€.
8 344 affaires nouvelles ont été souscrites pour un montant d'APE de 8 millions d'euros.
Les assurés ont en moyenne 42 ans.
Le coût moyen d'un sinistre incapacité (cumul des règlements et des provisions) est de 59 281€ pour une IJ moyenne de 112€ et une durée moyenne de 7,7 mois. L'incapable a en moyenne 45 ans.
La charge des prestations est de 76 millions d'euros (dont 60 millions pour l'incapacité/invalidité).

3. Comparaison du risque incapacité/invalidité entre notre portefeuille de TNS et notre portefeuille de salariés

La spécificité du risque incapacité/invalidité est qu'il est complexe à calibrer car il est à l'intersection d'une pathologie et de ses modalités de prise en charge (l'arrêt de travail est un état définit par une convention et non objectivement comme le décès). Le risque incapacité est sensible à l'évolution économique de la société, au comportement des assurés (l'aléa moral), ce qui en fait un risque volatile, et par conséquent mal connu. Un nombre important de paramètres affectent ce risque, ce qui en rend la tarification et le provisionnement très délicats.

Nous allons comparer quelques statistiques entre nos assurés TNS et nos assurés salariés ayant souscrit un contrat Generali en prévoyance individuelle.
Si la quasi-totalité des assurés se couvre en cas de décès, les salariés sont seulement 34% à se protéger contre l'incapacité et 25% contre le risque d'invalidité. La population des TNS se protège, elle, à 97% contre le risque d'incapacité et 92% pour l'invalidité. Cet écart s'explique par la différence d'indemnisation entre la sécurité sociale et le RSI comme vu précédemment.

L'âge moyen à la souscription est de 40 ans quel que soit la population.

Les hommes représentent 57% des salariés alors qu'ils sont quasiment 70% sur le portefeuille des TNS. Ce point reflète que la majorité des chefs d'entreprises, des artisans ou des professions libérales par exemples sont des hommes.

L'IJ moyenne souscrite est 25% plus importante pour les hommes que pour les femmes quel que soit la population. Cette différence notable peut s'expliquer par un écart de rémunération de l'ordre de 28% dans le secteur privé (chiffres 2010 de l'INSEE) entre les hommes et les femmes.

Nous remarquons également une différence importante entre le montant de l'IJ moyenne pour les TNS et les salariés : 123€ pour les TNS contre 38€ pour les salariés. Pour le montant mensuel de la rente invalidité, nous constatons le même phénomène : 850€ pour les salariés et 2 800€ pour les non-salariés. Ces écarts reflètent l'écart de rémunération et du régime de base entre les salariés et les TNS.

Les franchises souscrites en cas d'incapacité sont courtes pour la quasi-totalité des TNS (99% à moins de 30 jours) alors que les franchises courtes ne concernent que 70% des salariés.

Les garanties souscrites étant plus importantes pour les TNS, leurs primes le sont également : en moyenne 1 800€ de primes annuelles contre 400€ pour les salariés.

Concernant la sinistralité, en incapacité, l'âge moyen à la survenance est de 45 ans sur les deux portefeuilles alors que la durée moyenne en arrêt de travail est plus importante pour les salariés (8 mois) que pour les TNS (6 mois).

Sur l'invalidité, l'âge moyen à la survenance est de 48 ans pour les salariés contre 51 ans pour les TNS. La durée d'invalidité est en moyenne de 6 ans pour les salariés alors que les TNS restent 3 ans dans cet état.

Nous remarquons donc des différences de souscription et de sinistralité sur les deux populations étudiées concernant le risque incapacité/invalidité. Ce risque est donc sensible à la catégorie socio-professionnelle. En effet, les TNS se protègent plus contre ce risque : les montants des indemnités garanties sont beaucoup plus élevés (pour les TNS, l'IJ maximum possible est de 800€ et la rente mensuelle d'invalidité de 25 000€ alors, que pour les salariés, l'IJ maximum est de 200€ et la rente d'invalidité 6 000€ au maximum), les franchises sont plus courtes pour être indemnisé rapidement. Lorsqu'ils sont sinistrés, ils reprennent plus vite leur activité pour perdre le moins d'argent possible; ils ont un comportement différent face à la perte de revenu. Le portefeuille étant composé d'une majorité d'hommes, la rémunération du chef de famille ne doit pas trop diminuer pour faire face aux charges du foyer.

La population des TNS est donc spécifique sur le risque d'incapacité/invalidité, ce qui valide la construction de tables d'expérience car les tables règlementaires (population de salariés) ne reflètent pas la réalité de ce type d'assurés.

II. MODELE A ETATS ET QDD CHEZ GENERALI

Dans le cadre de Solvabilité 2, Generali a choisi d'estimer son risque par un modèle interne. Ainsi, dans ce paragraphe, nous exposons succinctement les grands principes du modèle à états pour la modélisation de la prévoyance individuelle et nous verrons également les procédures mises en place pour fiabiliser ce modèle à travers la Qualité Des Données (que nous appellerons par la suite QDD).

A. PRESENTATION DU MODELE A ETATS

Le modèle à états est basé sur des flux probabilisés, à partir de « model points », c'est-à-dire de sous-portefeuilles ayant des caractéristiques homogènes (âge à l'adhésion, ancienneté, garanties, franchises …). En effet, pour réduire les temps de calcul, la modélisation ne peut être faite contrat par contrat mais de manière agrégée en regroupant les contrats même si un certain « biais » peut être introduit.
Les effectifs et postes du compte de résultat sont projetés mensuellement jusqu'à un horizon de projection défini (dans le cadre du Best-Estimate, au maximum un an de projection en application du principe de la frontière des contrats car nos produits sont renouvelables par tacites reconduction ; dans le cadre de la valeur la projection d'effectue jusqu'à extinction du portefeuille).

Les résultats des simulations donnent plusieurs informations :

- flux et provisions pour chaque pas de projection
- comptes de résultat projetés pour l'ensemble du portefeuille, pour chaque année de projection
- Best Estimate : à la date de simulation il correspond à la somme des flux probables actualisés

La modélisation est effectuée à partir des données du portefeuille provenant du système de gestion à la date de simulation, ayant fait l'objet d'un traitement préalable avant intégration dans l'outil Prophet.

Trois populations sont créées pour chacun des produits :

- population des valides : les assurés valides ne percevant pas de prestation à la date de simulation
- population des incapables : les assurés percevant une indemnité journalière en cas d'incapacité à la date de simulation
- population des invalides : les assurés percevant une rente d'invalidité au titre d'une garantie invalidité à la date de simulation

Pour l'ensemble de ces populations, voici les variables principales nécessaires à la projection :

- l'état de l'assuré à la date de simulation: valide, incapable, invalide
- l'âge de l'assuré
- la durée courue depuis la souscription du contrat
- les indicatrices de souscription pour toutes les garanties modélisées : capital décès, rentes décès, rentes invalidité, indemnités journalières, exonération...
- la date d'entrée en incapacité pour les incapables et la date d'entrée en invalidité pour les invalides
- la franchise et la durée d'incapacité souscrites
- les montants garantis : montant du capital décès, montant de l'IJ, montant de la rente invalidité…
- le montant des primes par garanties

En plus des paramètres généraux ou spécifiques par produits (par exemples les taux d'actualisation, les âges maximum de couverture des garanties, taux de commissions, taux de frais, taux de réassurance…), le modèle nécessite l'utilisation de lois spécifiques au portefeuille pour déterminer le Best Estimate :

- la loi d'entrée en incapacité des valides
- la loi de maintien en incapacité
- la loi de passage d'incapacité en invalidité
- la loi de maintien en invalidité
- la loi de chute des contrats
- la loi de mortalité

Les constructions de ces lois, hors mortalité, sont exposées dans cette étude. Ces lois sont utilisées soit pour le calcul des provisions en Best Estimate, soit pour définir les effectifs des assurés dans chaque état pour chaque période de projection, soit pour projeter les cash-flows, soit pour obtenir les différents chocs à appliquer au modèle pour le calcul du capital économique (ce capital correspond au Solvency Capital Requirement représentant le capital cible nécessaire pour absorber le choc provoqué par des événements imprévus).

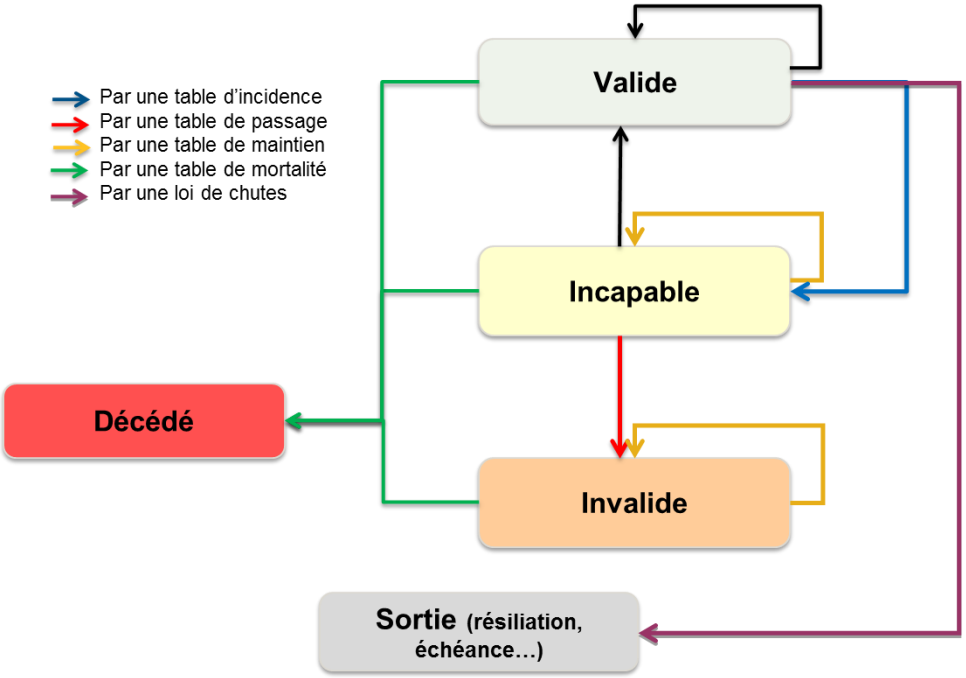
Le modèle est basé sur l'évolution de plusieurs états pour un assuré à chaque pas de projection:

- valide
- incapable
- invalide
- décédé
- sortis (résiliation ou échéance)

Par hypothèses, les effectifs d'assurés évoluent selon les principes suivants à chaque pas de projection :

- un assuré valide peut rester dans cet état, devenir incapable, décéder ou sortir du portefeuille (résiliation du contrat, échéance du contrat)
- un assuré incapable peut rester dans cet état, devenir invalide, décéder ou redevenir valide
- un assuré invalide peut rester dans cet état ou décéder. Nous faisons l'hypothèse qu'une personne invalide ne peut redevenir ni valide ni incapable
- un assuré qui a résilié son adhésion ou qui a son contrat arrivé à échéance ou qui est décédé n'est plus comptabilisé

Nous pouvons schématiser cette modélisation pour un assuré sur une période de projection par le graphe suivant :



Chaque flèche correspond à un passage d'un état à l'autre avec une probabilité définit par les différentes lois énoncées précédemment.

Le risque lié aux garanties incapacité et invalidité correspond essentiellement au risque de souscription vie, qui sont principalement des risques de morbidité dans la directive Solvabilité 2. Ce risque peut dévier sur deux composantes :

- le risque que davantage d'assurés entrent en incapacité/invalidité (risque d'incidence) donc le nombre d'assurés indemnisé augmentent
- le risque que les sinistrés se rétablissent moins qu'attendu (risque de maintien) donc les prestations durent plus longtemps

De plus, si la garantie exonération est souscrite, le montant des primes encaissées diminue car il y a plus de sinistrés sur une période plus longue qui ne paient pas leurs cotisations.

Les tables d'expériences construites dans ce mémoire vont permettre également d'évaluer l'impact des taux d'évolution de morbidité sur les engagements de l'assureur. Les chocs forfaitaires fixés par les spécifications de l'EIOPA sont :

- pour le risque de morbidité : majoration des taux d'incidence de 35% la 1ère année puis 25% les années suivantes
- pour le risque de longévité : abattement de 20% sur le taux de décès
- pour le risque de mortalité : hausse de 15% du taux de décès
- pour le risque de chute : hausse et baisse des taux de chutes de 20%…

Ces chocs impactent la projection des flux de primes et de sinistres pour le calcul du SCR. Ils permettent de prendre en compte les risques suivants:

- le risque d'estimation: les lois d'expérience sont des approximations des «vraies lois» sous-jacentes
- le risque d'échantillonnage: la charge sinistre est distribuée normalement autour de son espérance
- le risque de modèle: le modèle retenu constitue une approximation plus ou moins grossière de la réalité et ne reflète pas la réalité à 100%

A noter, que Solvabilité 2 exige la mise en place de processus et de procédures internes pour garantir la qualité des données alimentant le modèle. En effet, les données constituent l'information de base à partir de laquelle les calculs économiques et techniques sont effectués. Pour cette raison, la qualité des données a un impact majeur sur les calculs en affectant directement la cohérence et la précision des résultats. L'évaluation de la qualité des données repose sur la pertinence, l'exhaustivité et l'exactitude. Il faut également documenter la traçabilité de la donnée de la source jusqu'au modèle en décrivant les retraitements effectués.

B. LA QUALITE DES DONNEES CHEZ GENERALI

Dans la suite de ce mémoire, nous exposons les traitements et contrôles des données effectués pour nos études ; notions cruciales pour la directive Solvabilité 2 dans le cadre de l'évaluation des risques, de la fiabilité des résultats et des reportings. En effet, une QDD insuffisante pourrait motiver la constitution d'un capital add-on, fausser un calcul de Best Estimate ou de provisions techniques ou encore invalider l'utilisation du modèle interne (étant donné l'utilisation des statistiques du portefeuille pour paramétrer et alimenter celui-ci). Ce sujet est mentionné a de nombreuses reprises dans la directive, par exemples:

- A l'article 48 « apprécier la suffisance et la QDD utilisées dans le calcul des provisions techniques» , qui est une des mission de la fonction actuarielle

- A l'article 82, « mettre en place des processus et procédures internes de nature à garantir le caractère approprié, l'exhaustivité et l'exactitude des données utilisées dans le calcul de leurs provisions techniques»

Nous allons voir à présent, comment Generali a traité ce thème devenu incontournable dans la règlementation actuelle.

Depuis 2011, Generali France a formalisé son dispositif de gestion des risques et de contrôle interne en ligne avec la directive Solvabilité 2. Notamment pour la QDD, qui, à partir d'une méthodologie détaillée et définie en lien avec le Groupe, s'est progressivement déployée avec la mise en place d'une gouvernance dédiée, de processus et procédures précises.

La gouvernance de la QDD revet une dimension transverse qui impacte l'ensemble des activités de la société : gestionnaires, informaticiens, actuaires... Le rôle de cette instance est de valider le périmètre de la QDD, de valider les définitions des données, de piloter, de prioriser et de suivre les travaux liés à la QDD en prenant en compte les coûts et les enjeux mais aussi de définir les exigences en lien avec les règles métier ou avec la règlementation, d'effectuer des contrôles sur les données importantes, d'alerter et de diagnostiquer les anomalies éventuelles, de proposer des remédiations aussi bien sur le stock que sur les affaires nouvelles.

Un référentiel QDD a été mis en place pour partager, suivre et stocker tous les documents justifiant le tracé des données/fichiers, la formalisation et la documentation des contrôles sur les données. Voici les différents documents instaurés et présents dans cet outil :

- Un dictionnaire de données qui recense toutes les données nécessaires à l'alimentation du modèle interne avec des informations telles que la description, les synonymes (tous les métiers n'utilisent pas un même nom pour une même donnée), le format, des exemples de valeurs possibles, les systèmes de gestion concernés, les personnes responsables coté métier et coté informatique…

- Un diagramme des flux où chaque flux de données est tracé du système applicatif source jusqu'au modèle interne, en identifiant toutes les transformations de la donnée. Cette cartographie est composée d'une représentation graphique de l'ensemble des flux, d'un schéma de l'architecture technique ainsi que de détails sur les flux (description, fréquence, mode de mise à jour, contrôles de cohérence effectués, etc…)

- Un référentiel comprenant 16 contrôles types sur l'exhaustivité, l'exactitude et la pertinence. Les vérifications portent aussi bien sur le respect du planning et du format pour la livraison de la donnée que sur la cohérence de la donnée au sein du flux de données lui-

même, de son historique ou à l'aide d'un rapprochement avec un autre système, vérifier les relations entre des données corrélées…

- Une matrice de contrôles qui croise les 16 contrôles clés avec les données du dictionnaire et qui recense, pour chaque couple donnée / contrôle type, les contrôles informatiques ou métier existants ou à mettre en place

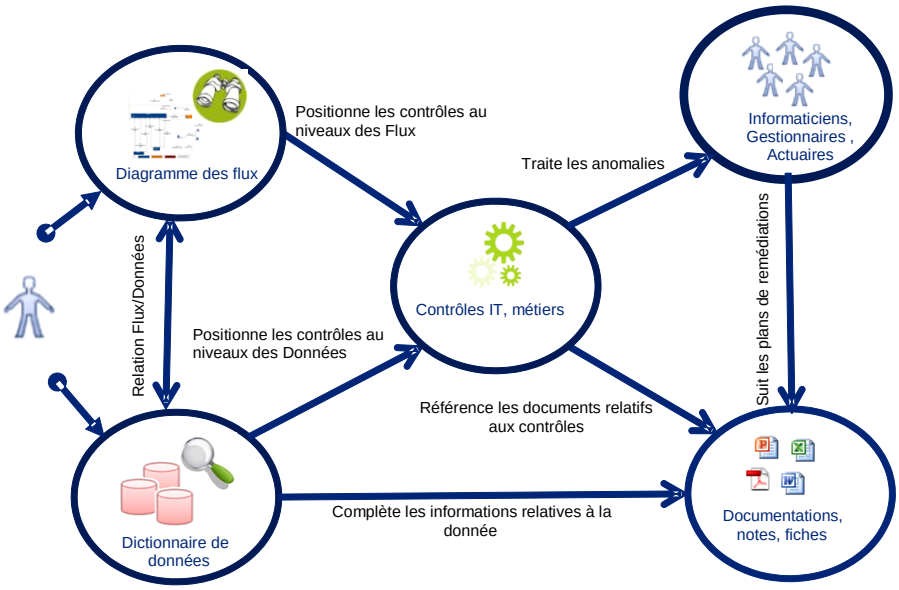
- Des fiches de contrôles par donnée du dictionnaire où tous les contrôles associés sont identifiés et détaillés : le critère QDD couvert (exhaustivité, exactitude ou pertinence), le type (contrôle technique, contrôle de valeur, contrôle de cohérence…), la nature (manuelle, automatique ou semi-automatique), la fréquence du contrôle, la localisation du contrôle, la significativité (aucun écart, anomalie < 1%...) et le plan de remédiation s'il existe

- Des indicateurs de qualité KPI (Key Performance Indicator) permettant d'évaluer le risque : pour chaque contrôle, sa qualité, son efficacité, son automatisation, son exécution… sont mesurés

- Pour le traitement des anomalies rencontrées, un service dédié est chargé de les suivre et de les faire corriger à travers des plans de remédiations impactant aussi bien les services de gestion que l'informatique. Des points sont organisés régulièrement avec les différents acteurs pour, par exemple, corriger des données directement dans le système de gestion mais aussi mettre en place des blocages informatiques pour pérenniser la remédiation sur les affaires nouvelles ou les nouveaux sinistres

Les différents contrôles mis en place dans ce cadre ont permis de relever à ce jour un certain nombre d'anomalies qui ont été corrigées ou qui sont en cours d'analyse et de correction : transformation ou historisation informatiques erronées, erreurs ou absence de saisies… Ces contrôles ont donc mis en évidence des anomalies liées au risque opérationnel.

Cette méthodologie QDD peut être schématisée de la façon suivante :



L'enjeu est de comprendre le cheminement des données clés et de positionner des contrôles à chaque étape du processus de transformation.

Pour satisfaire aux exigences Solvabilité 2 et au vue du volume de données de plus en plus important géré par les nombreux systèmes de gestion, Generali a lancé un projet d'infocentre unique. Il a pour objectif de permettre :

- L'utilisation d'une source de données unique et exhaustive quel que soit la branche considérée, ce qui permet d'homogénéiser et donc de renforcer la QDD et de faciliter la traçabilité des données
- La standardisation et l'automatisation de certains travaux qui permettent de réduire au maximum le risque opérationnel et d'accroitre la fiabilité des données afin de permettre aux actuaires de maximiser les phases d'analyse
- La sécurisation des flux permettant une optimisation de la QDD (définition, exhaustivité et cohérence, traçabilité, contrôles)

Solvabilité 2 va ainsi nous permettre d'améliorer la qualité de nos données et donc impacter directement la qualité de notre modèle interne. Nos calculs et nos risques sont ainsi plus précis, ce qui a également un impact direct sur la rentabilité de nos opérations.

La problématique QDD se renforce au sein de l'entreprise et un plan de communication se met actuellement en place pour sensibiliser tous les collaborateurs du groupe au sujet de la QDD : plus cette culture du risque est partagée et plus le risque pourra être maitrisé.

La QDD est utile dans le suivi des risques mais elle est aussi indispensable dans le monde digital et du Big Data qui se développent de plus en plus pour améliorer notre connaissance client au niveau de son comportement, de son mode de vie et de ses besoins par exemples. Les mégadonnées offrent une opportunité d'établir une tarification extrêmement fine, quasiment individualisée pour répondre à des clients de plus en plus désireux de payer le prix qui correspond à leur risque. Cela pourrait permettre également aux assureurs d'offrir de la prévention, des conseils ou un accompagnement personnalisé.

III. CONSTRUCTION DES TABLES DE MAINTIEN EN INCAPACITE, DE PASSAGE EN INVALIDITE ET DE MAINTIEN EN INVALIDITE

Les tables construites (maintien en incapacité, passage d'incapacité en invalidité, maintien en invalidité) vont nous permettre d'alimenter le modèle interne : estimation des provisions pour sinistres en cours Best Estimate, définition des probabilités de passage d'un état à l'autre, projection des cash-flows des sinistrés et calibrage des chocs pour le calcul du SCR.

Le but de cette partie est de détailler pas à pas l'obtention des tables pour les TNS couvert par un contrat de prévoyance.

Plusieurs étapes sont nécessaires pour construire chaque table :

- l'obtention de la base de données retraitées
- l'obtention des taux bruts
- le lissage de ces taux bruts

A. LES DONNEES

Nous construisons les tables d'expérience de morbidité sur le portefeuille contenant les produits PGM/PPL et ATOLL TNS. Par définition, une table d'expérience repose sur des données issues de l'observation d'un échantillon. Cette partie, détaillant les retraitements des données, a été la plus fastidieuse à réaliser, mais aussi la plus importante, car la qualité de l'étude dépend directement de la cohérence des données. Cette étape a été documentée pour lister, tracer et contrôler les données utilisées pour la construction des hypothèses alimentant le modèle dans le cadre de la qualité des données exigée par Solvabilité 2.

Les produits PGM/PLL et ATOLL ont changé de système de gestion en novembre 2008 : migration de tous les sinistres ouverts, des sinistres clos, des contrats en cours et des contrats sortis depuis le 01/01/2006 (soit une reprise d'historique de 3 ans). Une extraction de l'ancien système avait été faite en juin 2008 sur ces 2 produits pour la construction des tables d'expérience par un cabinet d'actuaires. Nous utiliserons donc ces extractions ainsi que les données actuelles disponibles dans l'applicatif existant pour récupérer la totalité des données sur ces produits. Ainsi pour avoir un historique complet sur ce portefeuille, nous allons travailler sur deux sources différentes, puis nous consoliderons ces bases pour obtenir une base unique.

1. Les bases de données de l'ancien système

Les données ont été extraites en juin 2008. L'historique des 2 produits peut être récupéré depuis leur début de commercialisation (PGM/PPL en 1989 et ATOLL en 2004).

a. Les fichiers sources

Nous disposons de trois fichiers de données au format texte :

- Sin PPL Atoll.txt : un fichier des sinistres ayant données lieu à une indemnisation jusqu'en juin 2008, ce fichier contient 12 961 lignes pour 12 921 couples contrat/sinistres. Les données utilisées sont :
    - Numéro de contrat
    - Numéro de sinistre
    - Code produit : pour distinguer le produit PPL/PGM du produit ATOLL
    - Garantie : pour distinguer l'incapacité de l'invalidité et pour supprimer les garanties qui ne nous intéressent pas dans cette étude telle que le décès, l'exonération…
    - Date de survenance du sinistre
    - Date de début de règlement
    - Date de fin de règlement
    - Montant du sinistre réglé sur la période
    - Montant garanti (définit à la souscription)

  Une agrégation des règlements par sinistres a déjà été effectuée lors de cette extraction par l'informatique.

- Prod PPL Atoll.txt et Prod PPL Atoll GV.txt : deux fichiers de données « Contrat » relatifs aux assurés ayant souscrit un contrat ATOLL ou PPL/PGM. Ces deux fichiers empilés comptent 336 727 lignes. Les données utilisées sont :
    - Numéro de contrat
    - Date de naissance de l'assuré
    - Code profession de l'assuré nous permettant de distinguer pour ATOLL les personnes salariées des travailleurs indépendants

b. Le traitement des données

Dans cette partie, nous détaillons tous les retraitements de données nécessaires à la construction d'une base ayant une qualité de données satisfaisantes comme exigée par Solvabilité 2 (traitements des valeurs manquantes, incorrectes ou aberrantes). Nous décrivons chaque étape de suppression ou de modification des données en expliquant la cause de ces manipulations et leurs conséquences notamment en termes de quantité de données restantes.
Le fichier des sinistres contient 7 586 sinistres pour PGM/PPL et 5 335 pour ATOLL (Atoll Salariés et Atoll TNS).

i. Numéros de contrat

Nous croisons les tables contrats et sinistres pour récupérer, dans notre base sinistre, les données relatives à l'assuré.
Pour pouvoir consolider cette base avec la base du nouveau système de gestion, nous devons faire le lien entre les anciens et les nouveaux numéros de contrats à l'aide d'une table de correspondance.
A noter qu'il n'existe pas de table de correspondance entre ancien numéro de sinistre et nouveau numéro de sinistre, ce qui va légèrement compliquer la consolidation de la base finale : nous devrons utiliser une clé permettant de lier les sinistres.

ii. Fiscalité

La fiscalité du contrat nous permet pour le produit ATOLL de distinguer les assurés salariés des assurés TNS. Or, cette variable n'existe pas dans l'extraction utilisée. Pour les contrats ayant migrés, nous récupérons l'information dans le système de gestion actuel. Pour les autres, nous utilisons le code profession pour créer une table de correspondance entre le code profession et la fiscalité. Lorsque la profession n'est pas renseignée, la fiscalité par défaut est TNS (concerne une dizaine de contrats).
Nous supprimons ainsi les sinistres ATOLL salariés pour ne conserver que les contrats TNS. Nous rappelons que l'intégralité des contrats PGM/PPL sont des contrats TNS.

iii. Garanties

Nous supprimons de la base sinistre toutes les garanties liées au décès toutes causes, décès accidentel, IAD, rentes éducations et rentes de conjoints.

Nous supprimons également toutes les garanties incapacité particulières telles que les frais professionnels, les IJ sur 1 an, les relais professionnels ainsi que l'exonération des cotisations, particularités qui pourraient biaiser notre étude : suppression de 2 545 lignes. En effet, l'exonération ne donne pas plus d'information que l'incapacité car, la plupart du temps, ces deux garanties sont couplées pour un même sinistre : l'une verse une indemnité journalière et l'autre rembourse le montant de la cotisation sur la même période d'arrêt de travail. Quant aux autres garanties, elles ne durent que 365 jours, en général, (l'observation de l'arrêt de travail est donc stoppée) et elles se cumulent avec une incapacité « classique » sur 3 ans. Elles n'apportent donc pas non plus une information supplémentaire et le sinistre à l'origine de cette indemnisation est déjà comptabilisé par l'incapacité « classique ».

iv. Sinistres particuliers

Nous supprimons également les sinistres sans règlement (identifiés par un montant de règlement à 0). En effet, ces sinistres n'ont pas été réglés parce qu'ils n'ont peut-être pas dépassé leur période de franchise ou parce que le dossier a été ouvert à tort ou toutes les pièces justificatives n'ont pas été réceptionnées …

v. Données manquantes ou aberrantes

Pour les données essentielles à l'étude, nous vérifions qu'il n'y ait pas de données manquantes ou aberrantes pour avoir une qualité de données satisfaisante. Nous faisons des contrôles sur la donnée elle-même mais également sur sa cohérence avec un autre flux de donnée.

Voici un tableau récapitulatif des contrôles et retraitements effectués sur cette base:

Nom de la donnée	Contrôle effectué	Volume concerné	Action
Date de naissance	Absence de donnée	0,0%	Récupération dans le système de gestion actuel
	= 1er du mois ou fin de mois	10,5%	Pas d'anomalie
Age à la survenance	< 18 ans ou > 66 ans	0,1%	Suppression
Date de survenance	Absence de donnée	-	Pas d'anomalie
	= 1er janvier	0,3%	Pas d'anomalie
	= 1er du mois	3,8%	Pas d'anomalie
Période d'indemnisation	Absence de donnée	-	Pas d'anomalie
Date de début d'indemnisation	< date de survenance	0,1%	Date de début d'indemnisation = Date de survenance
Date de fin d'indemnisation	< date de survenance	0,2%	Date de fin d'indemnisation = Date de début d'indemnisation + Nombre de jours indemnisés
Franchise* à 0	Valeur abs (IJ recalculée** - IJ garantie) > 5%	0,4%	Suppression

* Franchise = Date de début d'indemnisation - Date de survenance

** IJ recalculé = montant de prestation réglé / nombre de jours indemnisés

A noter que, pour un sinistre incapacité passant en invalidité, la date de survenance est identique entre les 2 garanties (elle correspond donc à la date d'entrée en incapacité). Nous faisons donc l'hypothèse que la date de survenance de l'invalidité correspond à la date de début d'indemnisation du sinistre invalidité.

A noter également que 62% des sinistres ont une franchise de 3 jours, 16% une franchise à 0 jour, 14% une franchise de 30 jours, 7% une franchise à 15 jours.

vi. Indicateur de passage

Pour les sinistres invalidité, nous créons un indicateur pour distinguer les passages d'incapacité en invalidité et les entrées directes en invalidité.

vii. Clé utilisée

Etant donné que notre base finale doit contenir une ligne pour un même sinistre, nous caractérisons les différents sinistres par une clé constituée de la concaténation des données suivantes :

- le numéro de contrat
- la date de survenance
- le type de garantie : incapacité ou invalidité

Le but étant que chaque clé corresponde à un seul sinistre.

viii. Synthèse des retraitements

Les différents traitements présentés dans les parties précédentes modifient les volumes de données disponibles. Il est important de mesurer la perte suite aux corrections apportées.

Voici le tableau récapitulatif des différentes suppressions de lignes effectuées :

	Nombre de lignes	
	Avant traitement	Retenues
Supression des contrats ATOLL salariés	12 961	12 046
Supression des garanties non concernées par l'étude	12 046	9 125
Supression des sinistres sans règlement	9 125	8 823
Suppression des dates de naissance aberrantes	8 823	8 813
Suppression des franchises aberrantes	8 813	8 776

Au final, voici le nombre de clés, donc de sinistres, pour cette première base TNS:

	ATOLL	PGMPPL	Total
Incap	2 653	5 679	8 332
Inval	9	111	120
Total	2 662	5 790	8 452

2. Les bases de données du système de gestion actuel

Les données sont disponibles mensuellement sous des tables SAS. Celles utilisées dans cette étude sont alimentées à fin juin 2013.

a. Les tables sources

Nous décrivons succinctement les variables utiles dans chaque table SAS.

i. Sinistres

La table sinistre correspond à l'ensemble des règlements effectués sur l'ensemble des produits et garanties. La profondeur d'historique des règlements est variable suivant les produits à cause des différentes migrations. A noter que, pour les contrats migrés de notre étude, une reprise d'historique de 3 ans a été faite. Nous allons donc avoir des doublons sur certains règlements lors de la consolidation des deux bases traitées.
Les données utilisées sont :

- Numéro de contrat
- Numéro de sinistre
- Code produit : pour distinguer le produit PPL/PGM du produit ATOLL
- Garantie : pour distinguer l'incapacité de l'invalidité et pour supprimer les garanties qui ne nous intéressent pas dans cette étude telle que le décès, l'exonération…
- Date de survenance du sinistre
- Date d'évènement (permet surtout de récupérer la date de passage d'incapacité en invalidité)
- Date de début de règlement
- Date de fin de règlement
- Montant du sinistre réglé sur la période
- Montant garanti (définit à la souscription)
- Situation du sinistre à la date d'extraction (sinistre ouvert ou sinistre clos)

Au total, cette base contient 853 788 lignes pour 198 334 couples numéro de contrat/numéro de sinistre.

ii. Contrats

La base contrat contient toutes les données relatives aux contrats. Elle nous permet surtout de récupérer le code fiscalité (distinction entre ATOLL TNS et ATOLL salariés).

iii. Personnes

Nous utilisons la base personne pour récupérer la date de naissance de l'assuré sinistré.

b. Le traitement des données

i. Produits

Nous ne conservons que les produits PGM/PPL et ATOLL.

ii. Garanties

Nous supprimons de la base sinistre toutes les garanties liées au décès toutes causes, décès accidentel, IAD, rentes éducations et rentes de conjoints.

Nous supprimons également toutes les garanties incapacité particulières telles que les frais professionnels, les relais professionnels, les IJ sur un an d'indemnisation, les IJ hospitalisation ainsi que l'exonération, particularités qui pourraient biaisées notre étude (comme vu précédemment dans le paragraphe III.A.1.b.iii).

iii. Sinistres particuliers

Nous supprimons les sinistres sans règlement (identifiés par un montant de règlement à 0) et les règlements annulés.

iv. Fiscalité

La fiscalité du contrat nous permet pour le produit ATOLL de distinguer les assurés salariés des assurés TNS. Pour la construction de notre table TNS nous supprimons donc les sinistres ATOLL salariés.

v. Données manquantes ou aberrantes

Pour les données essentielles à l'étude, nous vérifions qu'il n'y ait pas de données manquantes ou aberrantes pour avoir une qualité de données satisfaisante. Nous faisons des contrôles sur la donnée elle-même mais également sur sa cohérence avec un autre flux de donnée.
Voici un tableau récapitulatif des contrôles et éventuellement des retraitements effectués :

Nom de la donnée	Contrôle effectué	Volume concerné	Action
Date de naissance	Absence de donnée	-	Pas d'anomalie
	1er du mois ou fin de mois	8,00%	Pas d'anomalie
Age à la survenance	< 18 ans ou > 66 ans	0,00%	Suppression
Date de survenance	Absence de donnée	-	Pas d'anomalie
	1er janvier	0,40%	Pas d'anomalie
	1er du mois	4,70%	Pas d'anomalie
Période d'indemnisation	Absence de donnée	-	Pas d'anomalie
	Date début d'indemnisation ≤ Date fin d'indemnisation	-	Pas d'anomalie

Pour notre étude, il est cependant intéressant de s'attarder sur 2 données importantes où nous avons rencontré de nombreuses anomalies :

Dates de survenance

Ce système de gestion gère deux dates de survenance : la date de survenance et la date d'évènement différenciées en théorie par les règles suivantes :

- pour un sinistre en incapacité passant en invalidité, la date de survenance représente la date d'entrée en incapacité et la date d'évènement correspond à la date de passage en invalidité
- pour un sinistre incapacité sans passage en invalidité et sans autre particularité, les deux dates sont identiques
- pour un sinistre incapacité sans passage en invalidité mais avec une rechute ou un passage en incapacité partielle, la date de survenance correspond à la date initiale de l'arrêt et la date d'évènement correspond à la date de la rechute ou du passage de totale en partielle
- pour un sinistre en invalidité, la date de survenance est identique à la date d'évènement

Dans la réalité, ces règles ne sont pas toujours appliquées par les gestionnaires. Ainsi, nous observons beaucoup d'incohérences entre ces 2 dates. Cette problématique n'existait pas dans la première base, car le système de gestion ne gérait qu'une seule date. De plus, pour les sinistres issus de la migration, les dates de survenance d'incapacité et d'entrée en invalidité n'ont pas été correctement reprises. Nous devons donc, lors de la consolidation des 2 bases, faire un travail sur le retraitement de ces dates pour corriger au maximum les anomalies de gestion, d'une part, et les anomalies de migration, d'autre part.

Concernant les anomalies de gestions, nous avons identifié plusieurs cas :

- certaines rechutes ne sont pas considérées comme des rechutes mais des nouveaux sinistres sont créés sur le contrat avec une nouvelle date de survenance
- lors du passage en invalidité, le gestionnaire remplace la date de survenance de l'incapacité par la date d'entrée en invalidité
- lorsqu'un sinistre passe d'un état total à un état partiel, un nouveau sinistre est créé avec une nouvelle date de survenance

Ces points ont été remontés dans le cadre d'un chantier dédié à la qualité des données (comme exigé par Solvabilité 2):

- les services de gestion ont été sensibilisés sur l'impact que peut avoir une saisie incorrecte de ces données si importantes pour le calcul des provisions et l'analyse de la sinistralité
- un traitement manuel a été mis en place par les actuaires pour identifier ces sinistres et pour retrouver les bonnes dates de survenance pour la prise en compte du calcul des provisions
- régulièrement la liste de ces dates de survenance erronées sont transmises à la gestion pour correction dans le système

Nous modifions donc les dates de survenance identifiées comme erronées dans notre base pour améliorer la qualité de nos données. La prise en compte de cette liste modifie dans notre base 547 couples contrat/sinistre dont 318 déjà modifiés dans le système de gestion; nous corrigeons ainsi 229 dates de survenance.

Période d'indemnisation

En analysant nos sinistres, nous remarquons une anomalie sur les dates de début d'indemnisation : pour les sinistres migrés la date de début d'indemnisation correspond à la date de survenance et non à la date de début de règlement réelle. Lors de la consolidation des deux bases, nous devrons donc récupérer l'information correcte issue de l'ancien applicatif.

vi. Agrégation des lignes

La base des sinistres contient la totalité des règlements par contrat, sinistre et garantie. Or, nous n'avons pas besoin de ce niveau de détail pour notre étude. Nous agrégeons donc tous les règlements d'un même sinistre (couple contrat/sinistre) en retenant la date de début d'indemnisation minimum et la date de fin d'indemnisation maximum des différents règlements, ainsi que les autres informations communes à chaque règlement telles que la date de survenance, la date d'évènement, la garantie, l'âge à la survenance, le produit... Nous récupérons ainsi 22 770 lignes correspondant à 22 770 sinistres.

vii. Retraitements non effectuées

Etant donné les anomalies sur les dates de survenance et la date de début d'indemnisation, nous ferons le contrôle de cohérence sur les franchises, une fois ces dates retraitées donc après consolidation des deux bases.

viii. Synthèse des retraitements

Les différents traitements présentés dans les parties précédentes modifient les volumes de données disponibles. Il est important de mesurer la perte suite aux corrections apportées.
Voici le tableau récapitulatif des différentes suppressions/agrégations de lignes effectuées :

	Nombre de lignes	
	Avant traitement	Retenues
Supression des produits autres que PPL/PGM et ATOLL	853 788	182 880
Supression des garanties non concernées par l'étude	182 880	108 992
Supression des sinistres sans règlement	108 992	91 718
Supression des contrats ATOLL salariés	91 718	72 452
Suppression des dates de naissance aberrantes	72 452	72 451
Aggregation de tous les règlements d'un même sinistre	72 451	22 770

3. La base consolidée

La consolidation des deux bases sinistres provenant des deux systèmes de gestion présente diverses difficultés. En effet, elle ne se limite pas à une suppression des doublons. Pour pallier aux décisions prises lors de la migration des portefeuilles qui affectent la qualité des données, nous devons réaliser un grand nombre de retraitements :

- aucune correspondance entre ancien numéro de sinistre et nouveau numéro de sinistre : il faut donc utiliser une clé (concaténation du numéro de contrat et de la date de survenance) pour faire une correspondance entre un même arrêt de travail issu des deux bases

- la date de début d'indemnisation du sinistre a été reprise dans le nouvel applicatif par la date de survenance (c'est pourquoi nous obtenons quasiment 50% des franchises à 0 sur la seconde base sans retraitement). Lors de la consolidation, nous conservons donc la date de début d'indemnisation de l'ancien système lorsque le sinistre est présent dans les 2 bases

- pour les sinistres migrés, ni la date de passage en invalidité ni la date d'entrée en incapacité n'est correcte dans le système actuel; nous reprenons donc la date existante dans l'ancien applicatif

La concaténation des 2 bases crée une base de 31 222 lignes dont 30 661 pour l'incapacité et 561 lignes d'invalidité.

a. Le retraitement des données

i. Dates de naissance

Pour quelques contrats, nous avons une date de naissance différente entre les 2 systèmes. Nous décidons de conserver la date de naissance de l'applicatif actuel.

ii. Date de début d'indemnisation

Comme vu précédemment, la date de début d'indemnisation pour les sinistres migrés n'est pas correcte car elle correspond à la date de survenance. Nous retenons donc pour les sinistres présents dans les 2 sources la date de début d'indemnisation du système d'origine.

iii. Date de fin d'indemnisation

Pour la date de fin d'indemnisation, nous récupérons pour chaque clé (numéro de contrat/ date de survenance/ garantie) la date de fin d'indemnisation maximum quel que soit le système.

iv. Date d'entrée en incapacité

Cette date est l'une des variables les plus importantes pour la construction de la table de maintien en incapacité et de passage en invalidité, pouvant avoir un impact majeur sur les résultats. Cette donnée « sensible » n'est soumise à aucun contrôle ou blocage de l'applicatif, une liberté totale est donnée, en gestion, pour la saisie de cette date. Elle peut être interprétée de façon différente suivant le gestionnaire en charge du sinistre. Le retraitement lié à ce champ est donc l'un des plus compliqué. Nous définissons une règle générale permettant ainsi de récupérer un maximum de dates correctes en ne conservant qu'une seule date entre la date de survenance et la date d'évènement.

Plusieurs cas de retraitements ont été identifiés sur la base consolidée :

- Lorsque 2 sinistres d'un même contrat ont leurs périodes d'indemnisation qui se suivent alors on garde la date de survenance du 1er sinistre (soit 254 lignes concernées)
- Lorsque le sinistre n'existe que dans l'ancien système, il n'y a pas de problème car une seule date est renseignée : nous conservons la date de survenance (soit 2 004 lignes)
- Lorsque le sinistre est sur les 2 sources (soit 14 801 lignes):
    - Si les dates sont identiques (y compris la date de survenance est identique à la date d'évènement sur l'applicatif actuel) : nous gardons la date de survenance
    - Si les dates sont différentes : nous conservons la date du système d'origine
- Lorsque le sinistre n'est que sur le système de gestion actuel (soit 14 163 lignes) :
    - Si la date d'évènement et la date de survenance sont identiques : nous conservons la date d'évènement
    - Si les dates sont différentes :
        - Si la clé N°contrat/date de survenance/ garantie est unique, nous conservons la date d'évènement
        - Si la clé n'est pas unique, nous conservons la date de survenance minimum. En effet, si pour une date de survenance, nous avons plusieurs clés, cela signifie que plusieurs sinistres ont été créés avec une date d'évènement différente alors qu'ils se rattachent tous à un même sinistre

Cette étape ne supprime aucune ligne.

v. Indicateur de passage en invalidité

En consolidant les deux bases de données, nous pouvons identifier pour les invalides, ceux issus d'un passage d'incapacité en invalidité et ceux entrant directement en invalidité. Nous renseignons donc la variable indicateur de passage de la façon suivante pour un contrat ayant une garantie incapacité et invalidité sinistrées :

- si la date maximum de fin d'indemnisation en incapacité a moins de 31 jours d'écarts avec la date minimum de début d'indemnisation en invalidité alors nous considérons que c'est un passage en invalidité
- sinon l'invalidité est considérée comme directe

vi. Date d'entrée en invalidité et date de survenance d'incapacité liée

Une fois l'indicateur de passage renseigné, nous retraitons les dates d'entrée en invalidité pour les sinistres migrés. Lorsque le sinistre est présent sur les deux systèmes, nous décidons donc de conserver la date issue de l'applicatif d'origine.
Pour les entrées directes en invalidité, nous renseignons la date de survenance comme identique à la date d'entrée en invalidité.
Pour les passages, nous prenons comme date de survenance de l'incapacité, la date de survenance du dernier arrêt de travail qui précède le passage.

b. L'agrégation des lignes

L'agrégation des sinistres va nous permettre de supprimer les doublons entre les 2 systèmes de gestion mais aussi entre les différentes lignes de sinistres après tous les retraitements sur les dates effectués.
La clé utilisée pour l'agrégation est la concaténation du numéro de contrat, de la garantie et de la date de survenance retraitée.

Nous ne conservons alors que les données utiles à notre étude :

- Le numéro de contrat
- La garantie : incapacité ou invalidité
- Le nom du produit : ATOLL ou PGM/PPL
- La date de naissance retraitée de l'assuré
- La date de survenance en incapacité retraitée
- La date d'entrée en invalidité retraitée
- L'indicateur de passage en invalidité
- La date de début d'indemnisation retraitée minimum
- La date de fin d'indemnisation retraitée maximum
- La dernière situation du sinistre (clos, ouvert)

Cette étape nous permet d'obtenir une base de 24 503 lignes représentant 24 503 sinistres répartis de la manière suivante :

	ATOLL	PGMPPL	Total
Incap	12 537	11 541	24 078
Inval	130	295	425
Total	12 667	11 836	24 503

c. Le contrôle sur la qualité des données retraitées

Sur cette base consolidée et retraitée, nous effectuons quelques contrôles pour ne pas avoir d'incohérence et satisfaire les exigences de la directive.

i. Anomalie entre la date de survenance et la date de début d'indemnisation

Nous observons 10 sinistres pour lesquels nous avons une date de début d'indemnisation antérieure à la date de survenance. Nous modifions alors la date de début d'indemnisation de la façon suivante : Date de début d'indemnisation = Date de survenance

ii. Anomalie entre date d'entrée en invalidité et date de début d'indemnisation

De même, nous observons 2 sinistres pour lesquels nous avons une date de début d'indemnisation antérieure à la date d'entrée en invalidité. Nous modifions donc la date de début d'indemnisation de la façon suivante :

Date de début d'indemnisation = Date d'entrée en invalidité

iii. Anomalie entre la date de début et de fin d'indemnisation

La base contient 13 sinistres dont la date de début d'indemnisation est postérieure à la date de fin d'indemnisation. Après analyse de ces sinistres, la date de survenance et la date de début d'indemnisation semble correcte. La date de fin d'indemnisation est donc erronée. La date de fin nous servant pour l'identification des censures, nous ne modifierons pas ces dates; surtout que ces sinistres sont anciens et clos et ne sont donc pas censurés.
Nous conservons donc tous ces sinistres.

d. Le traitement des rechutes

Les rechutes ne sont pas toujours identifiées en tant que tel par les gestionnaires. Un nouveau numéro de sinistre est créé avec une nouvelle date de survenance. Pour retraiter cette anomalie, nous identifions les rechutes pour un même sinistre incapacité et les agrégeons pour n'avoir qu'un seul sinistre : nous considérons alors, comme défini dans les conditions générales, qu'il s'agit d'une rechute, si un contrat a deux sinistres dont la date de fin d'indemnisation du premier à moins de 45 jours d'écarts (pour ATOLL) ou moins de 60 jours d'écarts (pour PPL/PGM) avec la date de survenance du deuxième sinistre. Nous identifions ainsi 450 rechutes avec même parfois des rechutes multiples pour un même sinistre.
Lorsque deux sinistres (ou plus) sont regroupés, ils ne forment alors plus qu'un seul sinistre sur une seule ligne et les caractéristiques du nouveau sinistre sont définies de la façon suivante :

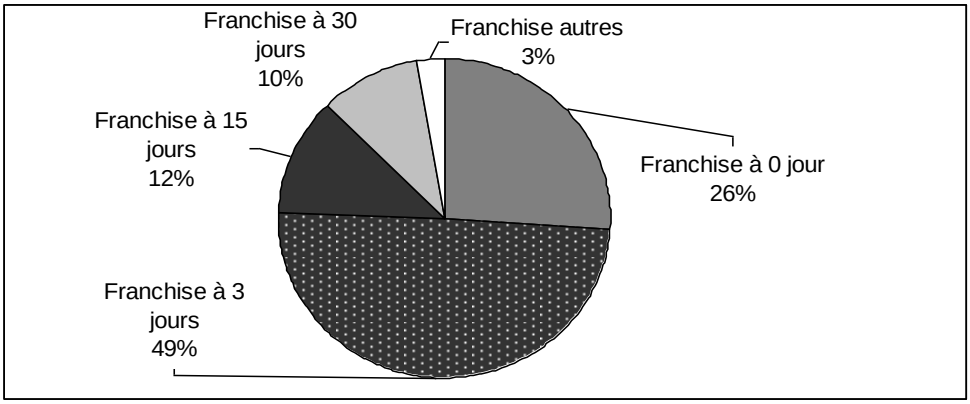
- sa date de survenance est la date de survenance minimum
- sa date de début d'indemnisation est égale à la date minimum
- sa date de fin d'indemnisation est égale à la date maximum

e. La franchise incapacité

Après retraitement des rechutes, nous calculons pour chaque sinistre le nombre de jours de la franchise de la façon suivante :

Franchise = Date de début d'indemnisation – Date de survenance

La répartition des franchises sur le portefeuille est la suivante :



Nous n'observons pas d'incohérence importante, les franchises les plus représentatives étant bien les franchises « officielles » des produits concernés. Il faudrait néanmoins vérifier les 26% de franchise à 0 jour. En effet, les franchises à 0 peuvent soit correspondre à des franchises résultants souvent d'un accident et parfois d'hospitalisation, à des franchises relatives, à des rechutes non identifiées créées sur un nouveau sinistre (ce cas vient d'être traité) ou à des anomalies d'erreur de saisie.

Nous faisons donc une étude spécifique sur les sinistres ayant une franchise à 0 jour pour détecter les anomalies et soit décider de les supprimer, soit de les corriger.

Pour cela, nous ne conservons que le premier règlement de chaque sinistre à étudier et sur cette base nous calculons :

- nombre de jours indemnisés = date de fin d'indemnisation – date de début d'indemnisation +1
- montant de l'IJ réglée = montant réglé / nombre de jours indemnisés
- montant de l'IJ garantie = montant garanti / 365 si le produit est PGM/PPL (produit en PASS)
  montant de l'IJ garantie = montant garanti si le produit est ATOLL (produit en euros)
- écart en pourcentage entre l'IJ garantie et l'IJ réglée

Si l'écart est inférieur à 10%, nous considérons qu'il n'y a pas d'anomalie et que la franchise est réellement à 0 jour. Sur les anomalies, nous testons si une franchise à 3 jours, 15 jours, 30 jours et 90 jours peut expliquer l'écart. Lorsque l'écart est expliqué nous définissons une nouvelle date de début d'indemnisation en l'ajustant à la franchise trouvée. Pour les sinistres restants, nous les considérons comme des anomalies et les supprimons de notre étude.

f. L'incapacité dépassant les 36 mois

Certains sinistres incapacité dépassent les 36 mois d'indemnisation. Plusieurs causes peuvent expliquer ce point :

- Dans certaines versions des Conditions Générales, la limite de durée d'indemnisation de l'incapacité n'est pas définie, ainsi le gestionnaire continue l'indemnisation en incapacité
- L'incapacité partielle prolonge de 6 mois la durée de l'indemnisation de l'incapacité totale
- Les rechutes peuvent décaler la date de fin d'indemnisation et donc dépasser les 36 mois
- Parfois, la gestion ne passe pas automatiquement un sinistre incapacité en invalidité (nous fournissons des listes de sinistres aux gestionnaires pour qu'ils les contrôlent et qu'ils les passent éventuellement en invalidité)

Les tables de maintien en incapacité n'excédant pas les 36 mois d'ancienneté, pour le calcul des provisions, nous forçons donc systématiquement les sinistrés en incapacité de plus de 36 mois en invalidité. Nous appliquons cette même méthode pour notre étude.

Après identification des cas concernés, nous gardons ces lignes d'incapacité sur 36 mois en modifiant la date de fin d'indemnisation par

$$\text{Date de fin d'indemnisation finale} = \text{Date de survenance} + 3 \text{ ans}$$

Nous doublons ainsi ces lignes pour avoir une ligne relative à l'incapacité et une ligne d'invalidité « fictive ». La date d'entrée en invalidité est ainsi définie par la date de fin d'indemnisation finale de l'incapacité + 1 jour.

Nous rajoutons ainsi 165 sinistres invalidité à notre base d'étude.

g. Le récapitulatif des retraitements

Les variables utilisées dans notre étude sont pour la plupart des données saisies ; pour un volume de données traité aussi important, elles peuvent donc potentiellement être incorrectes. Afin que notre étude soit la plus fiable possible, nous avons donc corrigé ces valeurs. Nous avons effectué différents tests pour détecter certaines incohérences, problématique augmentée par l'utilisation de 2 bases de données différentes issues de 2 systèmes de gestion différents ayant un fonctionnement différent, manipulant des variables parfois différentes; certains choix de migration retenus dégradant d'avantage la qualité de certaines données.

Selon l'anomalie constatée, nous avons choisi entre :

- Corriger automatiquement la valeur fausse ou manquante en la remplaçant par une valeur potentiellement correcte
- Procéder à une correction approximative à l'aide d'autres variables

- Agréger plusieurs sinistres en un seul
- Supprimer le sinistre contenant la valeur aberrante si une valeur correcte n'a pas pu être identifiée

Certaines valeurs étant liées entre elles, il convient de respecter un certain ordre lorsque nous procédons à la correction des données.
Voici la synthèse des suppressions et des retraitements effectués:

	Lignes supprimées	Lignes retraitées
Dates de naissance abérrantes	0,01%	
Franchises abérrantes	5,82%	
Fiscalités manquantes		0,08%
Période d'indemnisation < date de survenance		0,30%
Dates d'entrée en incapacité incorrectes		1,78%
Dates d'entrée en invalidité incorrectes		13,55%
Dates de début d'indemnisation incorrectes suite à la migration		15,44%

Le volume des données supprimées est faible car nous avons pu récupérer un certain nombre d'informations par des retraitements. L'anomalie identifiée la plus importante provient de la migration du système de gestion vers un autre applicatif qui peut être retraitée en croisant les données des deux systèmes. Cependant, nous pouvons considérer que la qualité des données est correcte au vu de ces différents pourcentages de données aberrantes et supprimées. En sensibilisant d'avantage les gestionnaires sur la saisie des différentes dates liées à un sinistre par exemple, la qualité des données pourrait s'améliorer et diminuer les retraitements manuels actuels pour le calcul des provisions et la ventilation par année de survenance des règlements. Les résultats des contrôles effectués, dans le cadre de la qualité des données, ne peuvent remettre en cause la validité des fichiers utilisés.

Voici le tableau récapitulatif des différentes suppressions/agrégations de lignes effectuées :

	Nombre de lignes	
	Avant traitement	Retenues
Aggregation des sinistres entre les deux bases	31 222	24 503
Traitement des rechutes	24 503	24 053
Suppression des franchises aberrantes	24 053	22 176
Ajout lié à l'incapacité supérieure à 36 mois	22 176	22 263

Notre base finale contient donc 22 263 enregistrements représentant chacun un sinistre. La ventilation de ces sinistres est la suivante :

	ATOLL	PGMPPL	Total
Sinistres incapacité	**10 491**	**11 260**	**21 751**
Sinistres invalidité	**157**	**355**	**512**
dont invalidités par passage	*138*	*304*	*442*
dont invalidités directes	*19*	*51*	*70*
Total	**10 648**	**11 615**	**22 263**

Sur l'incapacité, le volume des données parait suffisant, notamment pour une segmentation par âge sauf pour les âges extrêmes (nous verrons plus loin comment contourner ce problème). Rajouter une segmentation comme la catégorie socio-professionnelle ou la cause d'arrêt (maladie, hospitalisation, accident) réduirait la volumétrie des données observées sur ces axes et les lois construites ne seraient plus robustes.

L'ensemble des sinistres retenus nous servent pour :

- la construction de la table de maintien en incapacité pour tous les sinistres incapacité
- la construction de la table de passage d'incapacité en invalidité pour tous les sinistres incapacité et les sinistres invalidité issus du passage
- la construction de la table de maintien en invalidité pour tous les sinistres invalidité qu'ils soient directs ou issus du passage

Toute la difficulté de cette étape consiste à construire une base propre en supprimant le moins d'observations possibles pour conserver un volume de données suffisant et reflétant toutes les caractéristiques de notre portefeuille.

4. Les statistiques sur les données finales

Quel que soit la nature du sinistre, nous ne disposons pas du motif d'arrêt de la prestation, nous ne pourrons donc ni faire de statistiques sur cette variable ni s'en servi pour identifier les sinistres à censurer.

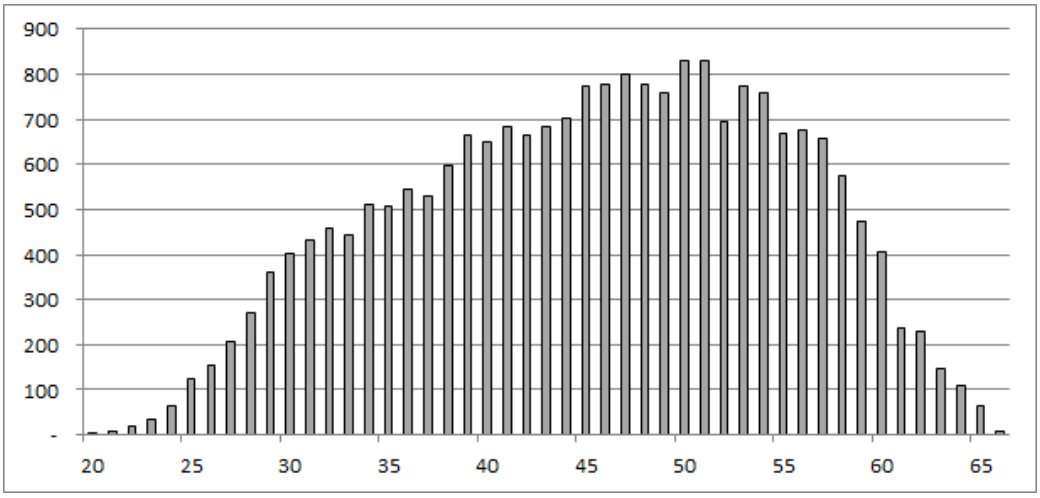
a. Les statistiques liées à l'incapacité

Notre base d'incapables TNS contient 21 751 sinistres.
Nous allons comparer les statistiques obtenues sur les TNS avec les statistiques de notre portefeuille de salariés.

i. L'âge à la survenance

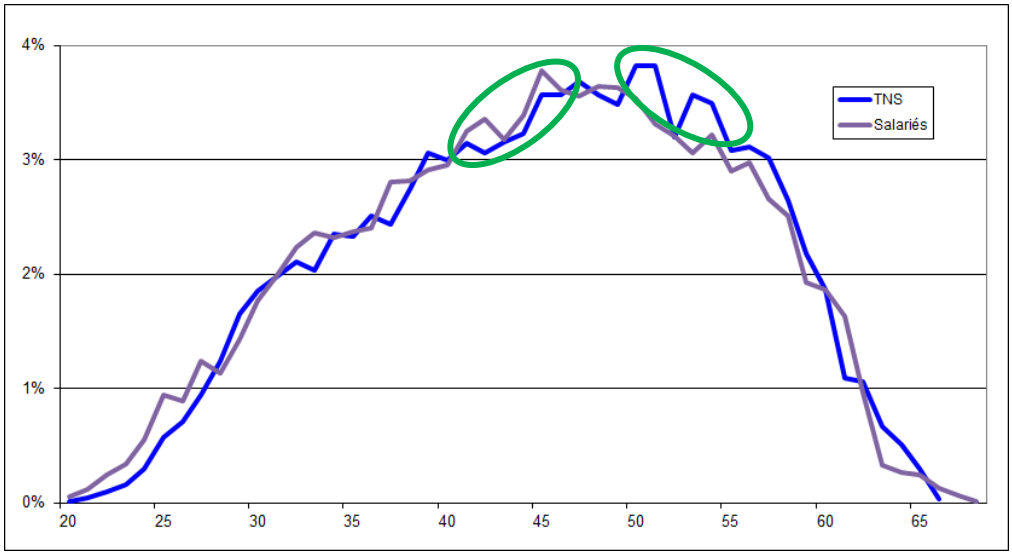
L'âge moyen à l'entrée en incapacité est de 45,3 ans, légèrement plus élevé que l'âge moyen du portefeuille (44,8 ans). Par contre, les salariés sont plus jeunes de 3 ans en moyenne lorsqu'ils entrent en arrêt de travail.

La répartition des incapables TNS suivant leur âge à la survenance est la suivante :



La table de maintien est construite par âge, il parait donc essentiel de pouvoir segmenter notre information selon l'âge. Cependant, nous observons que les âges extrêmes ne sont pas beaucoup représentés : moins de 200 sinistres par âge pour les jeunes de moins de 26 ans et les personnes âgées de plus de 63 ans. Néanmoins, cette difficulté peut être contournée : nous appliquerons la méthode qui consiste à calculer la loi pour chaque âge en y intégrant les données des âges environnants.

En comparant la pyramide des âges à la survenance entre notre portefeuille TNS et notre portefeuille salariés, nous observons que l'allure générale des courbes sont identiques. Nous pouvons noter néanmoins des pics vers 50 ans pour les TNS alors que pour les assurés salariés, l'entrée en incapacité est plus importante vers les 43 ans :

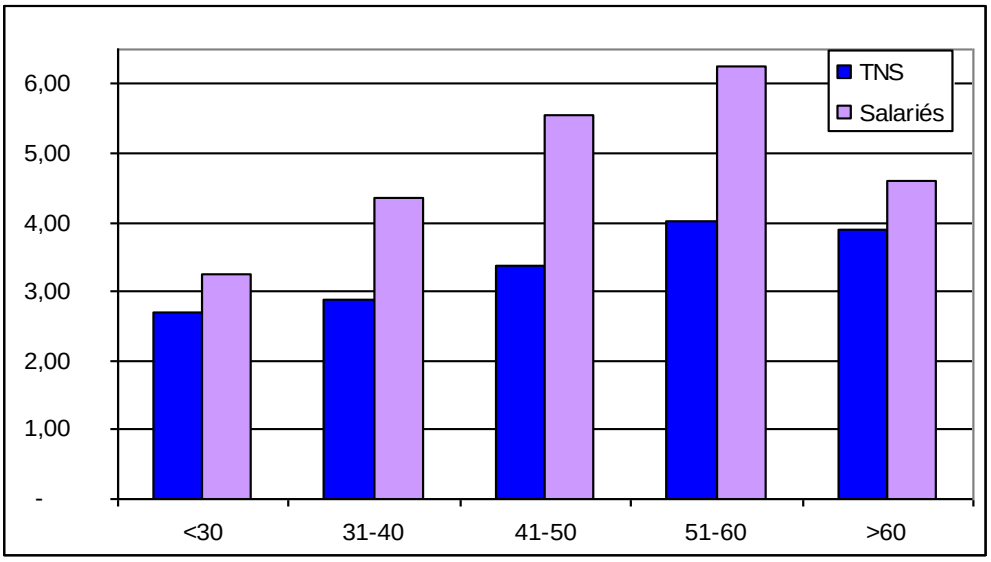


ii. La durée de l'arrêt de travail

La durée de l'arrêt de travail prend en compte la franchise et elle est calculée de la façon suivante :

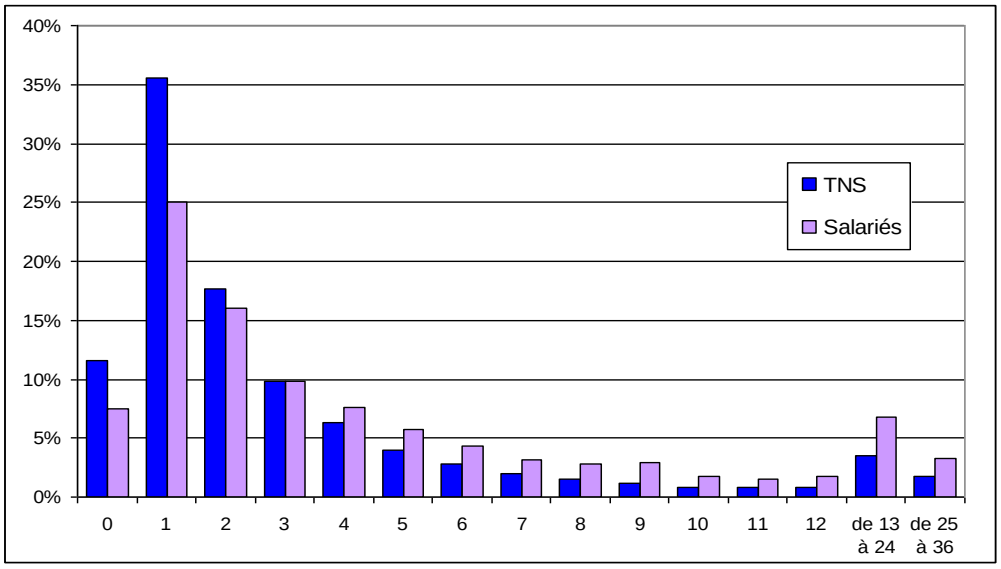
Durée de l'arrêt = Date de fin d'indemnisation – Date de survenance

La durée moyenne de l'arrêt en incapacité est de 3,4 mois pour les TNS contre 5,2 mois pour les salariés avec une répartition par tranche d'âge à la survenance suivante :



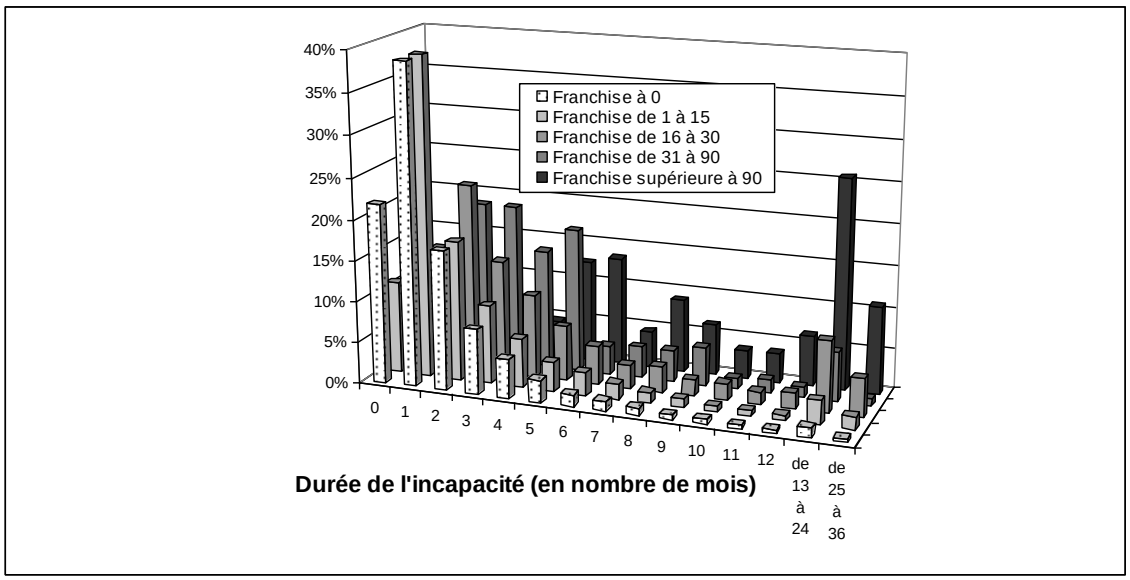
La durée de maintien en incapacité (en mois) dépend significativement de l'âge à la survenance quel que soit le portefeuille étudié. Plus la personne est âgée, plus elle met du temps à se rétablir en cas de sinistre. Nous remarquons toutefois une baisse pour les individus ayant plus de 60 ans sur les deux populations. Cette baisse peut être expliquée par les départs en retraite qui raccourcissent le maintien en incapacité. De plus, cette tranche est celle qui contient le moins d'observations (seulement 4% des sinistres).

Voici la répartition des sinistres suivant leur durée d'incapacité (en mois) :



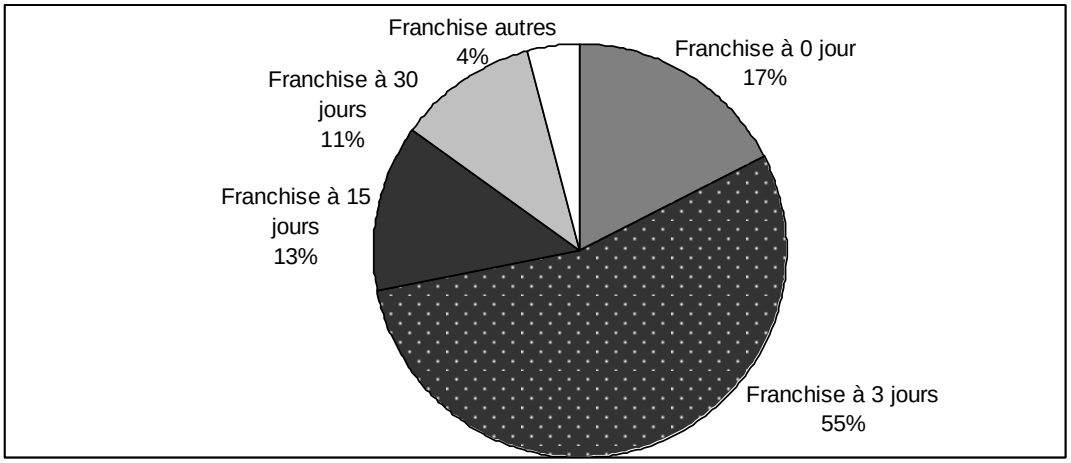
Nous constatons que, pour les TNS, les arrêts de travail sont courts : 75% durent moins de 3 mois, et, presque la moitié des sinistrés reprenne le travail au bout d'un mois, 5% des incapables seulement dépassant un an d'incapacité. Le portefeuille des salariés suit la même tendance. Cependant seul 60% des sinistres durent moins de 3 mois et 10% durent plus d'un an. Les TNS reprennent donc plus vite leur activité.

Lorsque nous observons, sur les TNS, la répartition des arrêts suivant leur durée et leur franchise, nous constatons également que plus la franchise augmente et plus l'arrêt dure longtemps car souvent les sinistres longs sont graves et couteux :



iii. La franchise

Voici la répartition des franchises des incapables de notre portefeuille TNS :

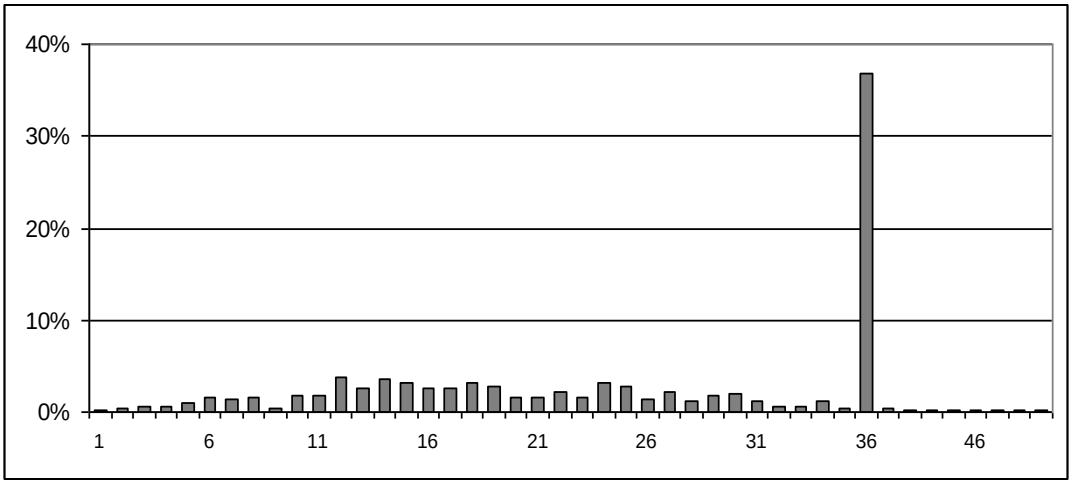


La franchise majoritaire est la franchise de 3 jours représentant plus de la moitié des sinistrés. A noter que, 96% des sinistres ont une franchise de moins de 30 jours donc une franchise courte, ce qui parait cohérent sur un portefeuille de TNS. La franchise à 0 jour correspond majoritairement à des sinistres causés par un accident.

Une étude a était faite sur la corrélation entre la sinistralité et la franchise souscrite. L'actuariat produit en charge de ce portefeuille a décidé d'appliquer des majorations significatives en 2015 pour les contrats ayant souscrits des franchises courtes. En effet, les sinistres ayant des franchises courtes coutent plus chers (car leurs fréquences sont importantes) à l'assureur qui porte donc un risque plus important.

iv. Le passage en invalidité

Nos observations contiennent 424 passages d'incapacité en invalidité.
Pour les TNS, seul 2% des incapables passent en invalidité au bout de 25 mois en moyenne alors que pour les 6% de passages des salariés, leurs passages s'effectuent vers les 20 mois d'ancienneté en moyenne.



La répartition de ces sinistres TNS suivant leur durée d'incapacité met bien en évidence le pic à 36 mois pour le passage en invalidité. Cette observation se retrouve sur le portefeuille des salariés. Nous observons également une dizaine d'incapables passant invalides après les 36 mois; dans ce cas, pour notre étude, nous les avons forcés à 36 mois d'ancienneté.
La volumétrie de l'exposition au risque (soit nos 21 751 incapables) est suffisante pour construire une table de passage en invalidité.

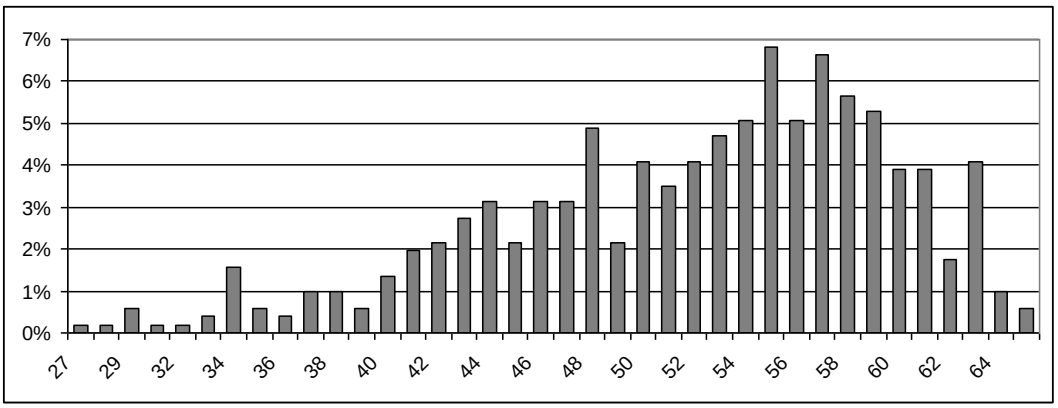
b. Les statistiques liées à l'invalidité

Notre base d'invalides contient seulement 512 sinistres invalidité.

i. L'âge à l'entrée en invalidité

L'âge moyen à l'entrée en invalidité est de 52 ans sans distinction particulière entre les invalidités de passage et les invalidités directes.

La répartition des invalides suivant leur âge à l'entrée est la suivante :



La répartition n'est pas très homogène ce qui s'explique par le faible nombre d'invalides observés : 512 sinistres.

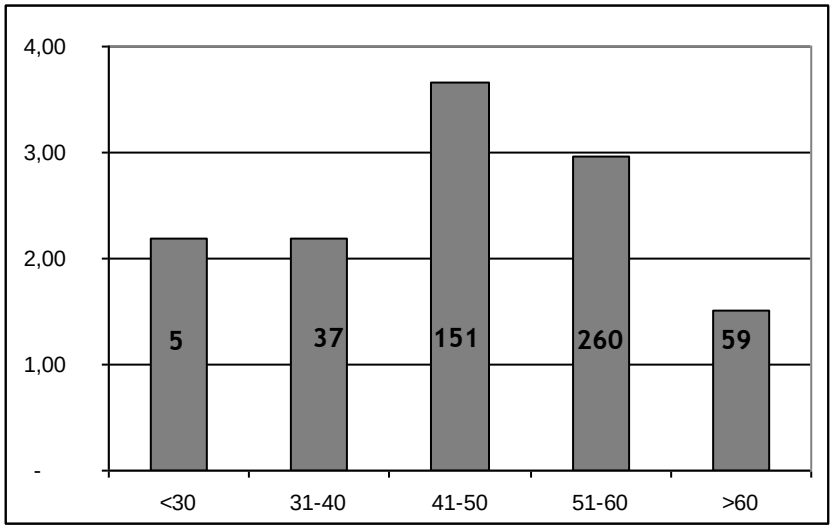
A noter que plus de 65% des sinistres ont plus de 50 ans à leur entrée en invalidité.

ii. La durée d'invalidité

La durée en invalidité est calculée de la façon suivante :

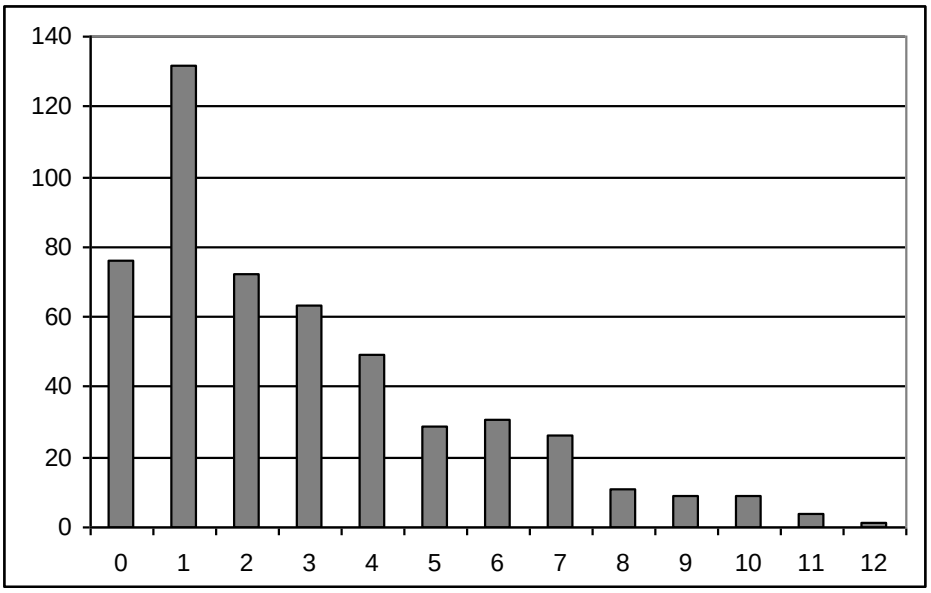
Durée de l'invalidité = Date de fin d'indemnisation – Date d'entrée en invalidité

La durée moyenne d'invalidité est de 2,9 années avec une répartition par tranche d'âge à l'entrée suivante :



Au vu du nombre de sinistres observés sur chaque tranche (chiffres en gras), nous ne pouvons rien déduire de fiable.

Voici la répartition des sinistres suivant leur durée d'invalidité (en années) :

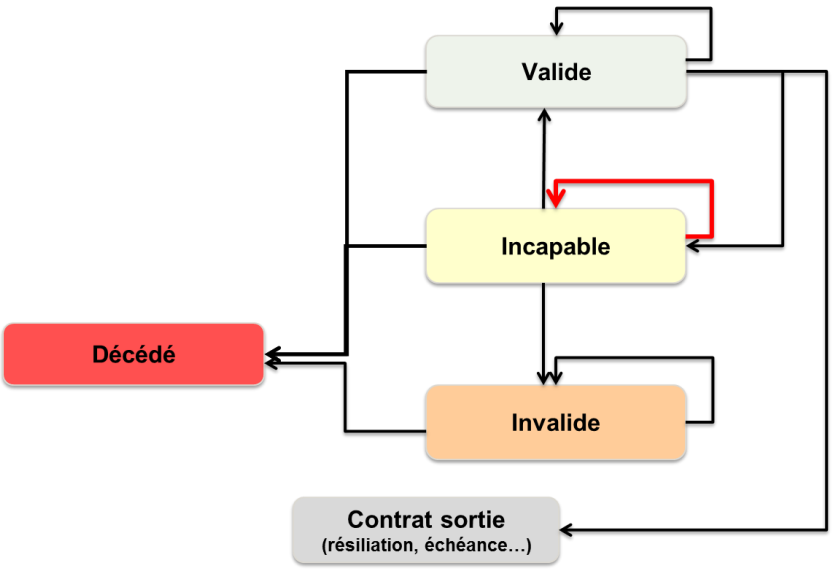


Nous constatons que 40% des sinistrés sont des invalidités courtes de 1 an ou moins. Dès les 4 ans d'invalidité dépassés, nous observons moins de 50 sinistres par ancienneté.

Au vue de ces statistiques, il parait impossible de construire une table de maintien en invalidité fiable, même par tranche d'âges et tranche d'anciennetés.

B. LA CONSTRUCTION DE LA TABLE DE MAINTIEN EN INCAPACITE

Pour le modèle à état, nous avons besoin d'une table de maintien en incapacité. En effet, elle est nécessaire pour la projection des cash-flows, pour le calcul des provisions incapacité en Best-Estimate et pour choquer le modèle. Elle sert également à faire évoluer l'état des assurés dans le modèle à chaque pas de projection avec la probabilité qu'un incapable reste dans cet état :



La table de maintien en incapacité se présente sous la forme d'un tableau à double entrée, avec en ligne, l'âge à l'entrée en incapacité, et en colonne, l'ancienneté en mois dans cet état, d'où l'importance de la qualité de ces deux données. Dans chaque cellule, figure $l_{x,t}$ c'est à dire le nombre de personnes encore en incapacité sur les 10 000 de référence présentes au départ (à l'ancienneté 0), pour l'âge à l'entrée x et une ancienneté de t mois.

1. Les données

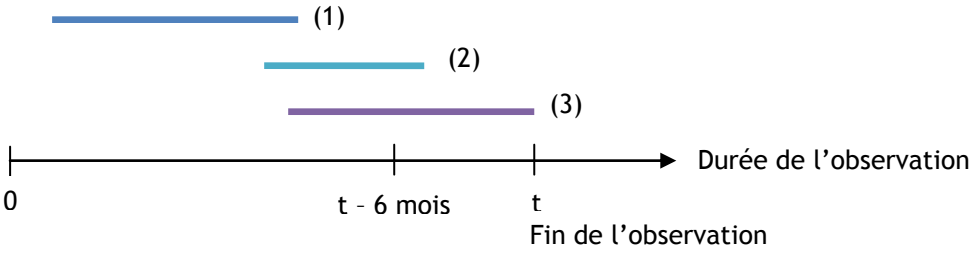
Estimer une loi par âge et par ancienneté nécessite d'observer les individus d'un échantillon sur une durée complète. Autant, nous avons pu récupérer tous les sinistres sur les 2 systèmes de gestion et nous avons donc bien le début d'observation pour tous les sinistres de notre étude (ayant dépassés leur franchise) autant nous avons arrêté d'observer la population à une date t où tous les sinistres n'étaient pas clos. Nous observons donc certains sinistres sur une durée incomplète. Il faut donc prendre en compte deux notions : les troncatures et les censures.

a. Les troncatures

Nous n'avons aucun sinistre en cours au début de notre observation étant donné que nous récupérons l'intégralité de l'historique de la sinistralité. Par contre, nos sinistres sont connus après leurs délais de franchise, nous n'observons donc aucun sinistre ne dépassant pas leur franchise. Tous les sinistres sont donc tronqués à gauche par leur durée de franchise.

b. Les censures

Nous considérons qu'un sinistre est censuré lorsqu'à la date d'observation t, le sinistre est toujours ouvert avec une date de fin d'indemnisation supérieure à la date d'observation – 6 mois. En effet, nous considérons que les sinistres ouverts avec une date de fin d'indemnisation de plus de 6 mois sont des anomalies de gestion et nous les considérons donc comme clos. Nous parlons de censures à droite :



Voici les différentes possibilités :

1 : Le sinistre est clos avant t - 6 : non censuré

1 : Le sinistre est ouvert avec une date de fin d'indemnisation < t – 6, nous le considérons comme clos : non censuré

2 : Le sinistre est clos entre t - 6 et t : non censuré

2 : Le sinistre est ouvert avec t - 6 < date de fin d'indemnisation < t : censuré

3 : Le sinistre est clos à la date t : non censuré

3 : Le sinistre est ouvert à la date t : censuré

Nos bases ne possédant pas de motif de sortie, nous ne pouvons pas censurer les sinistres à partir de ce critère. La clôture des sinistres correspond donc soit à un décès, soit à un passage en invalidité, soit à une reprise du travail.

c. La table finale des données

Chaque sinistre incapacité utilisé pour notre étude est caractérisé par les champs suivants :

- La clé du sinistre
- Le numéro de contrat
- La date de naissance de l'assuré
- La date et l'année de survenance du sinistre

- L'âge de l'assuré à la survenance
- Les dates de début et de fin d'indemnisation
- La durée de la franchise
- La durée de l'arrêt de travail (ancienneté)
- L'indicateur de censure

i. La clé du sinistre

La clé du sinistre est définie par la concaténation de 3 variables :

clé = Numéro de contrat + Date de survenance + Garantie

ii. L'âge à la survenance

La table de maintien se réalisant par âge à la survenance, nous arrondissons l'âge à l'entier le plus proche :

Age à la survenance = Arrondi (Date de survenance – Date de naissance)

iii. La durée de la franchise

La durée de la franchise est nécessaire pour l'établissement de cette loi car nous ne connaissons pas le maintien en incapacité des assurés avant cette franchise (notre étude est basée sur l'observation du sinistré une fois que la franchise est dépassée). Ainsi, tous les sinistres seront tronqués (censure à gauche) par leur durée de franchise. Cette durée est calculée en nombre de mois arrondi à l'entier le plu proche :

Durée de franchise = arrondi((Date de début d'indemnisation – Date de survenance) / 30,5)

iv. La durée de l'arrêt de travail

La durée de l'arrêt de travail est la caractéristique principale du sinistre. Elle est également calculée en nombre de mois arrondi :

Durée de l'arrêt = arrondi((Date de fin d'indemnisation – Date de survenance) / 30,5)

v. L'indicateur de censure

Pour chaque sinistre, nous définissons un indicateur de censure suivant la règle suivante :

Indicateur censure =

- 1 si le sinistre est ouvert avec une date de fin d'indemnisation > date d'observation – 6 mois
- 0 sinon

2. L'estimateur de Kaplan Meier

Pour pallier aux problèmes d'exploitation des données censurées ou tronquées pour l'estimation de notre loi de maintien, nous décidons d'utiliser l'estimateur non paramétrique du maximum de vraisemblance : l'estimateur de Kaplan Meier.

a. Notations

Voici les différentes notations utilisées dans la suite de l'étude :

- La durée de maintien en incapacité, c'est-à-dire le temps écoulé entre l'entrée et la sortie de l'état d'incapacité; nous la notons T

- La durée de maintien doit être discrétisée. Ainsi nous découpons T en m intervalles :

$[d_1 = 0, d_2 = 1[$; $[d_2 = 1, d_3 = 2[$; … ; $[d_{m-1}, d_m[$

Nous posons que T = d_i lorsque l'incapable sort de cet état dans l'intervalle $[d_i, d_{i+1}[$

- La probabilité que la durée de maintien en incapacité T dépasse la durée d_i est la fonction de survie S(d_i) :

$$S(d_i) = P(T \geq d_i)$$

Estimer la probabilité de survie revient à estimer la probabilité de la sortie d'incapacité

- Soit la probabilité q_i qu'un incapable sorte de cet état entre d_i et d_{i+1} sachant qu'il est encore en incapacité à l'instant d_i, c'est-à-dire la probabilité de sortie de l'état d'incapacité à la date d_i :

$$q_i = P(T = d_i | T \geq d_i) = \frac{P(d_i < T \leq d_{i+1})}{P(_i T \geq d_i)} = \frac{S(d_i) - S(d_{i+1})}{S(d_i)} = 1 - \frac{S(d_{i+1})}{S(d_i)}$$

Les q_i sont estimés par la méthode du maximum de vraisemblance à partir de nos observations constituant un échantillon de taille n.

Par définition, un individu a une probabilité q_i de sortir d'incapacité à la période d_i sachant qu'il était incapable à la période d_{i-1}. Par hypothèse, la probabilité de sortie d'incapacité pour chaque sinistré suit une loi de Bernoulli de paramètre q_i. Nous supposons que, la sortie de l'état d'incapacité de chaque individu est indépendante. Ainsi, le nombre de sorties s_i à la période d_i suit une loi binomiale de paramètres (n_i, q_i) où n_i représente l'effectif sous risque pour la période d_i.

- Soit p_i la probabilité qu'un incapable se maintienne dans cet état entre d_i et d_{i+1} sachant qu'il est encore en incapacité à l'instant d_i :

$$p_i = 1 - q_i = \frac{S(d_{i+1})}{S(d_i)} = P(T \geq d_{i+1} \mid T \geq d_i)$$

- Rappelons que, si X est une variable aléatoire suivant une loi binomiale de paramètre (n,p) alors $P(X = k) = C_n^k p^k (1-p)^{n-k}$

- Rappelons également que, la fonction de vraisemblance pour une variable aléatoire discrète X = (x_1 ,x_2 … x_n), un n échantillon indépendant et identiquement distribué, dont on veut estimer un paramètre θ, est définit par $L(x_1,...,x_n \mid \theta) = \prod_{i=1}^{n} P_\theta(X = x_i)$

b. Obtention de l'estimateur

Le principe de l'estimation de Kaplan-Meier appliqué au maintien en incapacité repose sur une idée simple : être encore dans l'état d'incapacité après un instant d_i, c'est être incapable juste avant d_i et ne pas sortir de l'état à cet instant.

En terme de probabilité, cette affirmation se définit par la relation suivante:

$$\begin{aligned} S(d_i) &= P(T \geq d_i) \\ &= P(T \geq d_i \mid T \geq d_{i-1}) * P(T \geq d_{i-1}) \\ &= P(T \geq d_i \mid T \geq d_{i-1}) * \ldots * P(T \geq d_2 \mid T \geq 0) * P(T \geq 0) \\ &= p_i * p_{i-1} * \cdots * p_1 \\ &= \prod_{j<i} (1 - q_j) \end{aligned}$$

La fonction de survie S(d_i) peut donc être estimée par le produit des estimations des q_i.

Compte tenu de la discrétisation de la variable T, il est nécessaire de fixer une convention sur l'ordre d'intervention des entrées, sorties et censures au cours d'une même période. Nous décidons que les entrées précèdent les sorties, qui précèdent elles-mêmes les censures (a priori la convention choisi concernant l'ordre d'intervention des grandeurs n'a pas d'impact significatif sur le maintien en incapacité).

L'effectif sous risque n_i se définit donc comme suit :

$$n_1 = N_0 + t_1$$

$$n_{i+1} = n_i - s_i - c_i + t_{i+1}$$

avec n_i l'effectif sous risque pour la période d_i

s_i le nombre de sorties non censurée pour la période d_i

c_i le nombre de censures à droite pour la période d_i

t_i le nombre de troncatures à gauche pour la période d_i

Sous la condition d'indépendance entre le processus de censure et celui régissant les sorties, les taux q_i ne sont pas modifiés et les estimateurs définis précédemment sont toujours valables.

La fonction de vraisemblance que nous cherchons à maximiser s'écrit :

$$\mathrm{L}=\prod_{\mathrm{i}=1}^{\mathrm{m-1}} C_{n_i}^{s_i} * q_i^{\,s_i} * \left(1-q_i\right)^{n_i - s_i}$$

En passant à la log-vraisemblance on obtient :

$$\ln L = \sum_{i=1}^{m-1} \left(C_{n_i}^{s_i} + s_i \ln q_i + \left(n_i - s_i\right) \ln\left(1-q_i\right) \right)$$

La maximisation de l'estimateur est effective lorsque les dérivées partielles du premier ordre s'annulent et lorsque les dérivées partielles secondes sont négatives. Ainsi nous devons résoudre $\forall q_i$:

$$\begin{cases} \dfrac{d \ln L}{dq_i} = \dfrac{s_i}{q_i} - \dfrac{\left(n_i - s_i\right)}{\left(1-q_i\right)} = 0 \\ \dfrac{\partial^2 \ln L}{\partial {q_i}^2} = -\dfrac{s_i}{{q_i}^2} - \dfrac{\left(n_i - s_i\right)}{\left(1-q_i\right)^2} \le 0 \end{cases}$$

La première égalité nous donne la formule des estimateurs des taux de sortie : $\hat{q}_i = \dfrac{s_i}{n_i}$

En remplaçant les q_i par leurs estimateurs $\hat{q}_i$ dans la formule de la fonction de survie déterminée précédemment nous obtenons l'estimateur de Kaplan Meier de la fonction de maintien en incapacité :

$$\hat{S}(d_i) = \prod_{j<i} (1-\hat{q}_j) = \prod_{j<i} (1-\frac{s_j}{n_j})$$

c. Application de l'estimateur

Pour appliquer cette méthode à nos observations, nous construisons un tableau par année de survenance, âge de survenance et période i contenant :

- t_i le nombre de troncatures à gauches c'est-à-dire le nombre d'entrée des arrêts de travail tronqué par leur durée de franchise égale à i
- c_i le nombre d'incapables censurés de durée i
- s_i le nombre de sorties d'arrêts de travail non censurés de durée i

Par exemple la ligne ci-dessous signifie que :

- 15 sinistres, d'année de survenance 2010 et d'âge de survenance 40 ans ont une durée de franchise d'1 mois et entrent donc en indemnisation à partir de cette période
- 1 seul sinistre, d'année de survenance 2010 et d'âge de survenance 40 ans est censuré au bout du premier mois
- 8 sinistres, d'année de survenance 2010 et d'âge de survenance 40 ans sortent au bout d'une durée d'un mois

Année de survenance	Age de survenance	Période	t_i	c_i	s_i
2010	40	1	15	1	8

Etant donné, la taille de notre échantillon, nous ne pouvons faire aucune segmentation de la table autre que l'âge et l'ancienneté.

Pour construire, par âge et ancienneté notre population sous risque n_i, et, pour avoir des statistiques plus robustes, nous décidons d'appliquer la méthode de glissements par âge pour pondérer les t_i, c_i et s_i. Cette méthode consiste à regrouper directement les âges de manière à satisfaire les contraintes de volume, tout en conservant un résultat propre à chaque âge. Nous décidons ainsi d'intégrer les données des âges adjacents, c'est-à-dire pour calculer un taux à l'âge x, nous sommons les données des âges x-1, x et x+1.

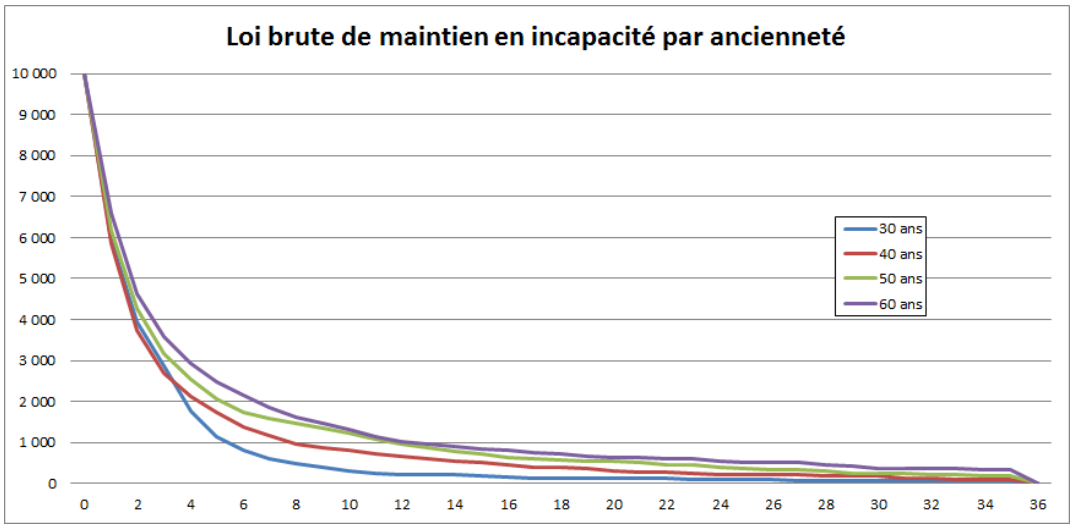
Cependant, même après cette pondération, pour les âges extrêmes, le nombre d'effectif pondéré sous risque reste faible. Pour avoir un volume significatif d'individus, nous regroupons les âges extrêmes pour avoir un minimum de 2 000 observations par âge. Nous regroupons ainsi les sinistres des 20 à 26 ans ainsi que de 64 à 66 ans. Ces deux paliers d'âge constitués « disparaitront » lors du lissage des taux.

3. L'obtention des taux bruts

Pour la construction de la table de maintien en incapacité, nous partons sur des données de 20 à 66 ans et de 0 à 36 mois avec un pas mensuel.

a. Résultat

Le graphe suivant représente les lois brutes de maintien en incapacité pour les âges de 30, 40, 50 et 60 ans.



Nous constatons que le niveau des courbes croit avec l'âge, ce qui confirme que l'âge est un facteur aggravant du maintien en incapacité. Que ce soit durant les 2 premiers mois ou à partir du 24ème mois les courbes sont très proches. Ainsi à partir de 2 ans, l'âge intervient moins dans le maintien en incapacité, seule la gravité du sinistre compte.

b. Intervalles de confiance de l'estimateur

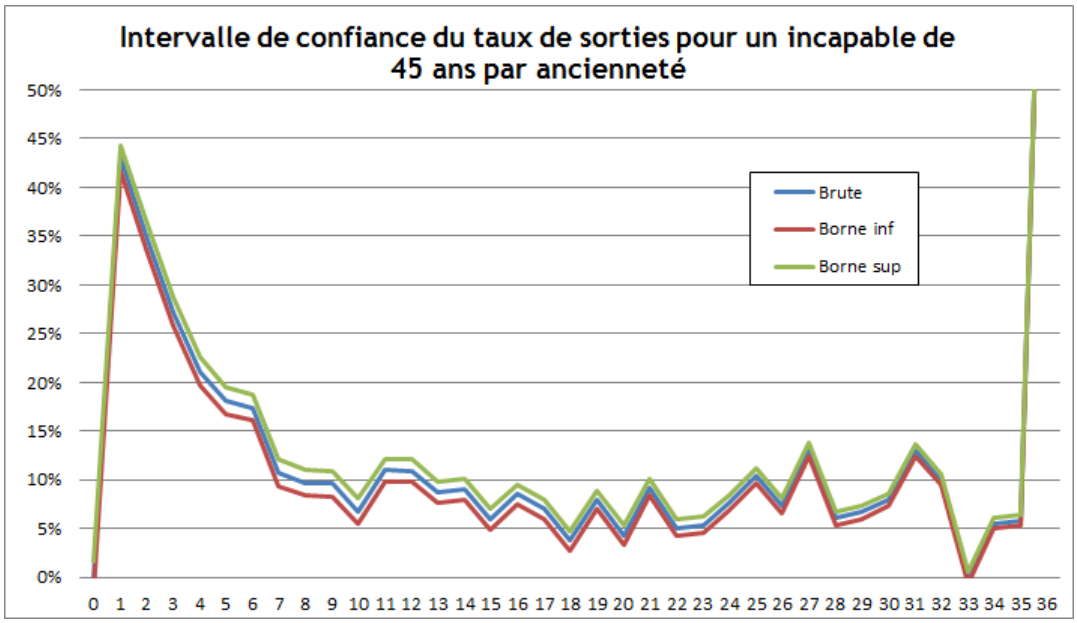
La fonction de survie étant une proportion, une approximation normale pour son intervalle de confiance peut être retenue.

En effet, $\dfrac{\sqrt{n}\left[\hat{S}(d_i) - S(d_i)\right]}{\sqrt{\overset{\wedge}{Var}(\hat{S}(d_i))}} \xrightarrow[n \longrightarrow +\infty]{L} \mathrm{N}(0,1)$.

Nous obtenons ainsi la formule suivante pour estimer l'intervalle de confiance à 95% :

$$IC_{95\%} = \left[\hat{S}(d_i) \pm 1{,}96\sqrt{Var(\hat{S}(d_i))}\right].$$

Nos assurés entrant en incapacité en moyenne à 45 ans, regardons l'intervalle de confiance à cet âge :



La largeur de l'intervalle de confiance varie selon l'ancienneté, néanmoins elle est plutôt étroite sur une grande partie du graphique; notre estimation est donc plutôt précise.
Nous constatons la même tendance sur d'autres âges (voir les graphiques en annexe 1).

4. L'obtention des taux lissés

La méthode décrite précédemment, nous permet d'obtenir une table de maintien en incapacité brute. De plus, pour palier à un volume de données faibles sur les âges extrêmes, nous les avons regroupés et nous avons ainsi formé 2 paliers. Les irrégularités de certains taux doivent être lissées pour éviter d'avoir des écarts trop importants entre deux anciennetés ou deux âges consécutifs. Le lissage appliqué doit permettre d'obtenir une loi plus régulière ainsi qu'une faible dispersion entre les taux bruts et les taux lissés.
Nous choisissons la méthode d'ajustement non paramétrique de Whittaker Henderson qui permet d'effectuer un double lissage suivant l'âge à l'entrée en incapacité et l'ancienneté. De plus, elle nous a paru la méthode la plus adaptée puisqu'ayant déjà fait ses preuves dans le monde de l'actuariat en ce qui concerne la construction de table.

a. Méthode du lissage de Whittaker Henderson

La méthode de Whittaker-Henderson combine un critère de fidélité et un critère de régularité en recherchant les valeurs ajustées minimisant la somme de ces deux critères. Il est en plus possible de jouer sur certains paramètres permettant de favoriser l'un ou l'autre de ces deux critères (voir la méthode détaillée en annexe 2).

Pour utiliser, la méthode de Whittaker-Henderson, il est nécessaire de choisir trois paramètres :

- Les poids w_i affectés au respect de chaque donnée

- Le paramètre z tel que z-1 est le degré du polynôme que nous considérons comme le standard de la régularité

- Le paramètre h d'arbitrage entre lissage et fidélité

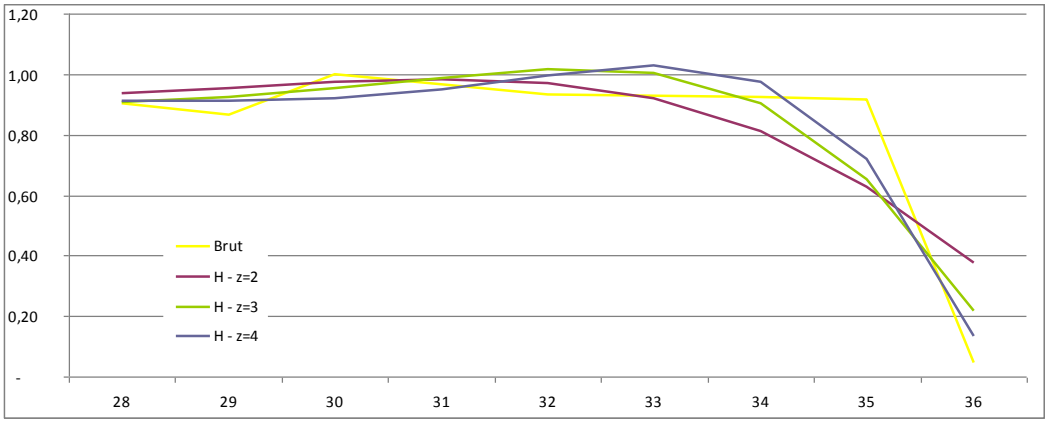
Nous choisissons de lisser les taux de maintien en incapacité. Pour cela, nous allons comparer sur des graphiques les différents lissages obtenus avec des paramètres différents, tout d'abord sur un lissage horizontal (selon l'ancienneté), puis sur un lissage vertical selon l'âge.

Pour le lissage horizontal :

Les graphiques ci-dessous correspondent à la courbe d'un incapable ayant 50 ans à la survenance.

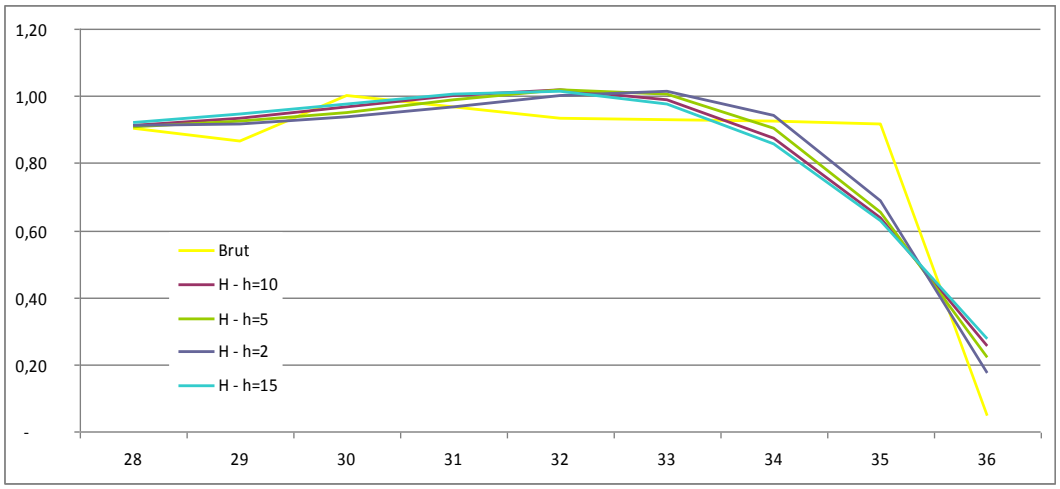
- Choix pour les poids w_i : nous décidons d'affecter un poids identique à l'ensemble de nos données donc w_i=1

- Choix du degré du polynôme z : nous devons choisir la valeur de z qui lisse le mieux la courbe tout en restant fidèle aux données brutes. Les courbes lissées sont très proches quel que soit la valeur de z de 0 à 28 mois. Nous zoomons donc de 28 à 36 mois :



Nous choisissons donc z=3. En effet, z=4 ne semble pas lisser suffisamment la courbe et au contraire z=2 la lisse trop.

- Choix du paramètre h : Le paramètre h est un paramètre d'arbitrage entre fidélité et régularité. Plus la valeur de h est importante et plus la courbe obtenue est régulière. De même que pour le lissage précédent, les courbes sont très proches pour les différents h pour les anciennetés de 0 à 28 mois. Nous ne regardons donc que les courbes pour les anciennetés supérieures à 28 mois :



Un bon compromis peut être h=5 car pour cette valeur de h, la courbe lissée se rapproche de la courbe brute tout en diminuant les trop grandes irrégularités.

Nous appliquons la même démarche pour déterminer les paramètres du lissage vertical en traçant les graphiques pour une ancienneté de 4 mois. Les paramètres retenus pour ce lissage sont les mêmes (w_i =1, z = 3 et h = 5).

b. Validation du lissage

La validité du lissage est testée à travers un test du Khi-2 qui compare les taux lissés aux taux bruts en calculant la statistique sur l'âge et l'ancienneté :

$$W = \sum_{x=\text{âgemin}}^{\text{âgemax}} \sum_{t=\text{ancienneté min}}^{\text{ancienneté max}} \frac{\left(D_{x,t} - \hat{D}_{x,t}\right)^2}{\hat{D}_{x,t}}$$

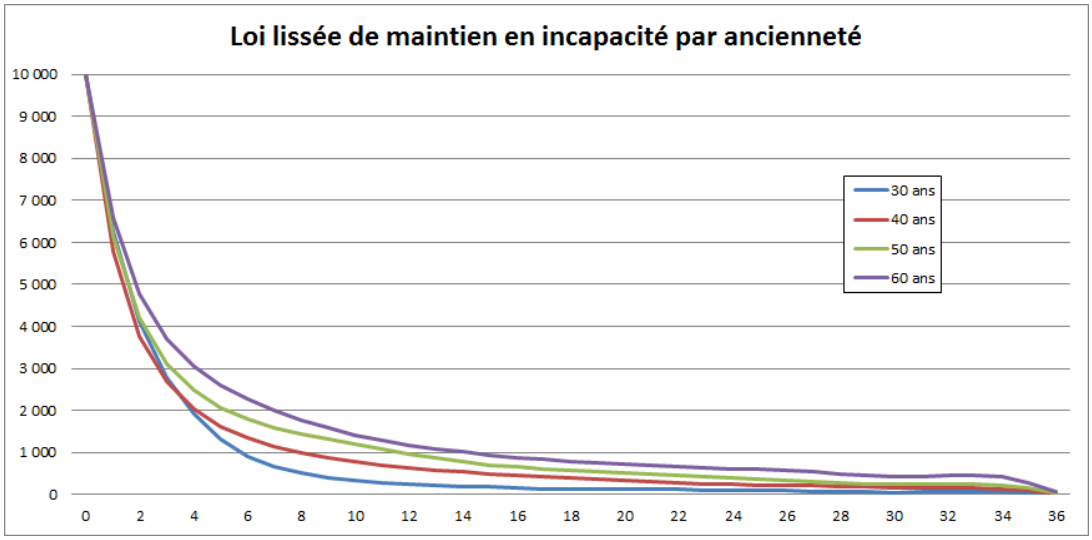
Avec $D_{x,t}$ le taux brut de maintien en incapacité pour un individu de x ans et une ancienneté de t mois et $\hat{D}_{x,t}$ le taux lissé correspondant.

Ainsi W suit asymptotiquement une loi du Khi-2 à 81 degrés de liberté (le degrés de libertés correspond au nombre d'occurrences c'est-à-dire le nombre d'ancienneté et le nombre d'âges de la table de maintien – 2).

En prenant un seuil de 95%, la région critique de notre étude est $\{W > 61\}$. La valeur de la statistique W sur notre table est de 23, ce qui nous conduit à accepter l'ajustement.

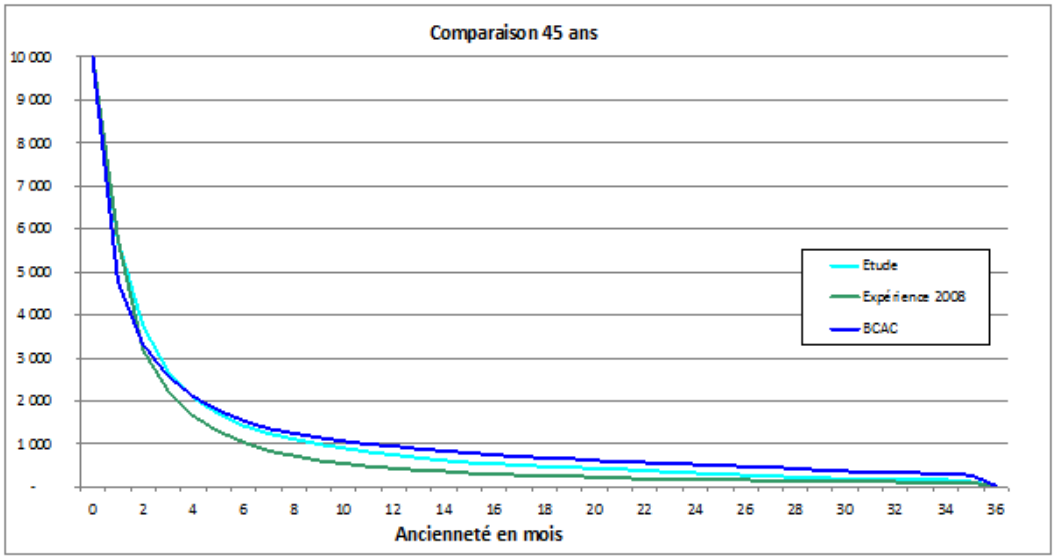
c. Résultat

Le graphique ci -dessous représente la loi lissée de maintien en incapacité à différents âges suivant l'ancienneté :



Comme pour les courbes brutes, nous observons que l'âge est un facteur aggravant du maintien en incapacité, même si pour les 3 premiers mois, la loi de maintien à 30 ans est au-dessus de celle à 40 ans, comportement certainement expliqué par le faible nombre d'observations à cet âge. Nous remarquons également que l'écart entre les courbes n'est pas constant ; il diminue de plus en plus à partir de 12 mois d'ancienneté jusqu'à se confondre au bout des 3 ans. En effet, au bout de quelques mois d'arrêt de travail, c'est d'avantage la gravité du sinistre qui détermine le maintien en incapacité, indépendamment de l'âge de l'assuré.

L'âge moyen d'entrée en incapacité de notre portefeuille est de 45,4 ans. Nous comparons donc la courbe lissée de maintien en incapacité obtenue à 45 ans avec celle du BCAC et avec l'ancienne table d'expérience :



Pour un incapable de 45 ans à la survenance, nous observons qu'à partir du 4ème mois, la loi de notre étude se situe entre la loi règlementaire du BCAC et la loi d'expérience construite il y a 5 ans. Les 3 premiers mois, notre loi se situe au-dessus des 2 autres donc une sinistralité plus importante.

En général, nous observons, quel que soit l'âge à la survenance (graphes présents en annexe 3) :

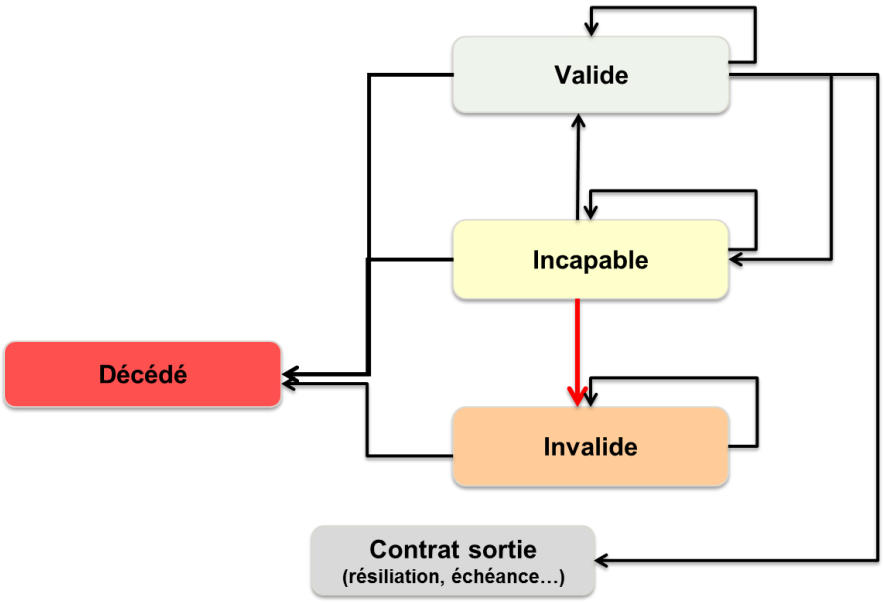
- les 3 premiers mois, notre loi est toujours au-dessus des deux autres
- sur les autres mois, notre maintien en incapacité est systématiquement supérieur à celui établi en 2008 mais inférieur à celui du BCAC

Concernant la comparaison avec la loi du BCAC, nous pouvons faire différentes classes d'âges :

- pour les moins de 25 ans, jusqu'à environ 2 ans d'ancienneté, nous observons une augmentation relativement importante surtout au cours des premiers mois
- entre 26 et 40 ans, nous obtenons un nombre d'incapables supérieur les premiers mois (entre 8 et 15 mois) puis notre loi passe en-dessous de celle du BCAC
- entre 41 et 60 ans, notre maintien est inférieur au maintien règlementaire pour l'ensemble des mois écoulés sauf sur les 3 premiers mois
- les plus de 60 ans, notre loi est au-dessus de celle du BCAC jusqu'à 9 mois puis elle est inférieure

C. LA CONSTRUCTION DE LA TABLE DE PASSAGE EN INVALIDITE

Pour le modèle interne, nous avons également besoin de la table de passage d'incapacité en invalidité pour projeter les flux, pour calculer les provisions en Best-Estimate, pour calibrer les différents chocs à appliquer au modèle sur le risque d'incidence et pour modéliser l'évolution de l'état d'un incapable :



La table de passage en invalidité (ou aussi appelé table de rente en attente) se présente sous la forme d'un tableau à double entrée, avec en ligne l'âge à l'entrée en incapacité, et en colonne l'ancienneté en mois dans cet état. Dans chaque cellule, figure $S_{x,n}$ le nombre de personnes passant en invalidité sur les 10 000 incapables du départ à l'âge d'entrée x en incapacité depuis n mois.

1. Les données

Dans notre base finale de données, nous avons construit un indicateur de passage pour identifier les invalides issus d'un passage d'incapacité et les invalides directs. Pour notre table de passage en invalidité, nous ne sélectionnons donc que les invalidités issues d'un passage soit 441 sinistres. Les invalidités directes servant seulement pour la construction de la loi de maintien en invalidité. Ce volume de données peut sembler faible, cependant, ce qui compte pour la construction de cette table c'est l'exposition au risque soit le nombre d'incapable, qui lui est suffisant.

a. Les variables utiles

Pour construire notre table de passages en invalidité, nous avons besoin des variables suivantes :

- la clé du sinistre (identique au maintien en incapacité)
- le numéro de contrat
- la date de naissance de l'assuré
- la date d'entrée en invalidité
- la date de survenance en incapacité
- l'âge arrondi de l'assuré à la survenance de l'incapacité (en année)
- l'ancienneté arrondie de l'incapacité au moment du passage en invalidité (en mois)

Dans ce cas, nous n'avons pas la notion de censure ou troncature car nous souhaitons juste savoir combien de sinistres incapacité sont passés en invalidité, nous n'avons pas besoin d'observer les passages sur la totalité de leur état.

b. Le tableau récapitulatif

Nous construisons ainsi un tableau comme pour l'incapacité par année de survenance, âge de survenance et période i (i représentant ici l'ancienneté en incapacité avant le passage en invalidité) contenant le nombre de passages observés.

Par exemple, la ligne ci-dessous signifie que 2 sinistres incapacité, d'année de survenance 2010 et d'âge de survenance 40 ans sont passés en invalidité au bout de 15 mois d'incapacité.

Année de survenance	Age de survenance	Période	Nombre
2010	40	15	2

c. Le volume des observations

Etant donné que pour la table de maintien en incapacité, nous n'avons pas pu faire de segmentation autre que l'âge et l'ancienneté, nous n'en faisons pas non plus sur la table de passage en invalidité.
De la même façon que pour l'incapacité, nous appliquons la méthode des glissements par âge pour pondérés le nombre de passages en invalidité : pour calculer un taux à l'âge x, nous sommons les données des âges x-1, x et x+1.
Pour rester cohérent avec le maintien en incapacité, nous regroupons les passages des âges extrêmes : de 20 à 26 ans ainsi que de 64 à 66 ans.

2. L'obtention des taux bruts

Pour la construction de la table de passage en invalidité, nous partons sur des données de 20 à 66 ans et une ancienneté de 1 à 36 mois sans autre segmentation.

Pour obtenir la table de passage en invalidité qui correspond aux nombres de passages, nous cherchons tout d'abord à estimer un taux de passage brut. Pour cela, nous récupérons la table construite précédemment pour le maintien en incapacité contenant les effectifs sous risque n_i. Nous avons les nombres de passages observés pondérés. Nous obtenons alors, pour chaque âge et chaque ancienneté, un taux de passage en invalidité au bout de i mois d'incapacité définit par

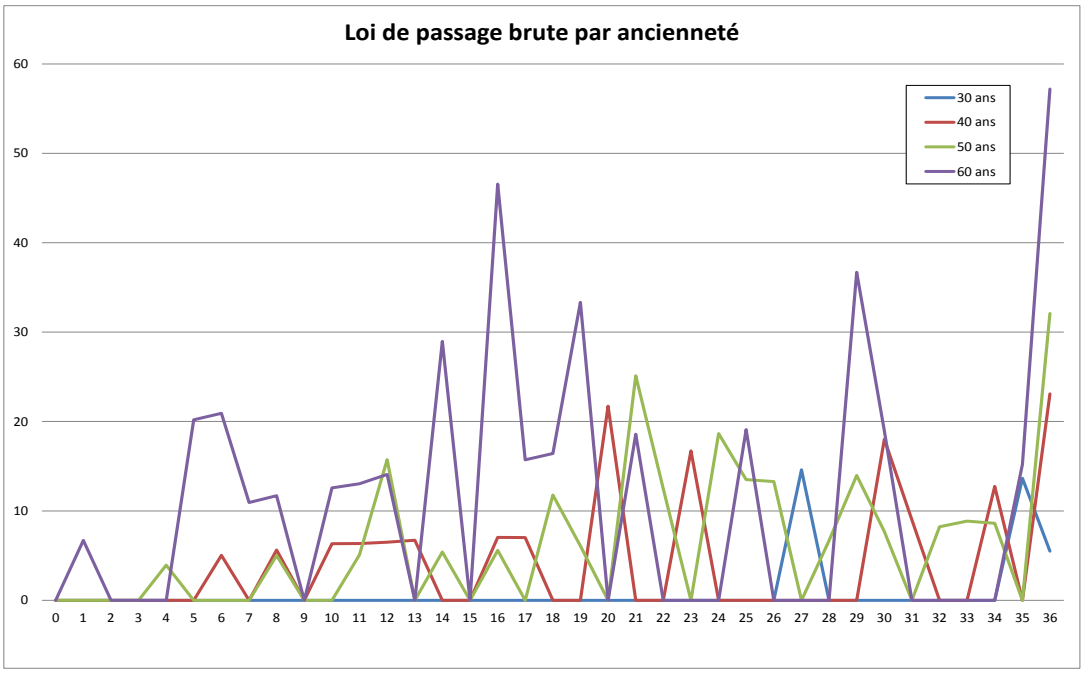
$$\text{Taux passage}_i = \frac{\text{Nb passage}_i}{n_i}$$

avec Nb passage_i : le nombre d'incapable passant en invalidité au bout de i mois d'incapacité

n_i : l'effectif sous risque en incapacité le mois i (définit dans la partie de la table de maintien en incapacité)

Pour obtenir la loi brute de passage par âge et par ancienneté (donc un nombre de passages), nous multiplions le taux de passage obtenu par le nombre d'incapables de la loi de maintien en incapacité déterminée précédemment.

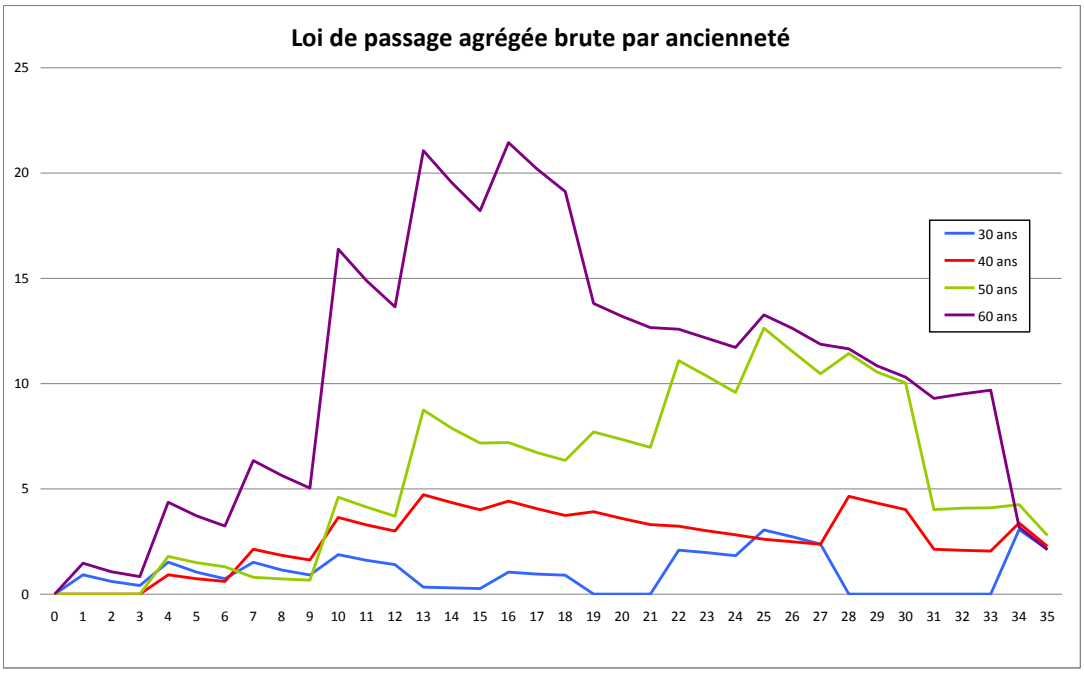
Le graphe suivant représente les nombres bruts de passage en invalidité pour les âges de 30, 40, 50 et 60 ans.



Nous constatons une volatilité importante suivant l'âge et l'ancienneté : la maille étant très fine (âge par âge, mois d'ancienneté par mois d'ancienneté), de nombreux taux sont égaux à zéro. Nous

décidons donc de faire des regroupements pour obtenir des lois plus robustes. L'approche choisie consiste à rendre la maille moins fine pour l'obtention du taux de passage : nous avons regroupé les passages et les expositions d'incapacité en 5 classes d'âge (moins de 35 ans, 36 - 40 ans, 41 - 45 ans, 46 - 50 ans et plus de 51 ans), et agrégé les anciennetés par trimestre. Ceci permet d'avoir des résultats plus robustes, et d'obtenir une structure par âge et par ancienneté mieux visible.
Par convention, nous posons que le taux de passage à 36 mois est de 100% pour tous âges, c'est-à-dire que nous définissons le nombre de passages en invalidité comme égale aux nombre d'incapables de la table de maintien du 36^{e} mois. Ainsi nous forçons tous les incapables en invalidité au bout des trois ans d'incapacité.

Une fois que le taux de passage agrégé est obtenu, nous le multiplions par les effectifs de la table de maintien en incapacité. Ainsi « les paliers » obtenus par agrégation disparaissent :



En général, nous constatons que le niveau des courbes croit avec l'âge, ce qui confirme que l'âge est également un facteur de passage en invalidité.

3. L'obtention des taux lissés

Comme pour la loi de maintien en incapacité, nous utilisons la même méthode de lissage sur le taux de passage : le lissage par la méthode de Whittaker-Henderson.

Le lissage appliqué doit permettre d'obtenir une loi plus régulière ainsi qu'une faible dispersion entre les taux bruts et les taux lissés.

a. Choix des paramètres pour le lissage de Whittaker Henderson

Comme précédemment, pour choisir les paramètres du lissage, nous comparons les graphiques obtenus avec les différents paramètres testés à la courbe brute.

Nous conservons toujours un poids identiques pour l'ensemble des données : w_i=1.

Nous retenons les paramètres suivants pour le lissage horizontal et vertical :

- degré du polynôme : z = 3

- paramètre : h = 5

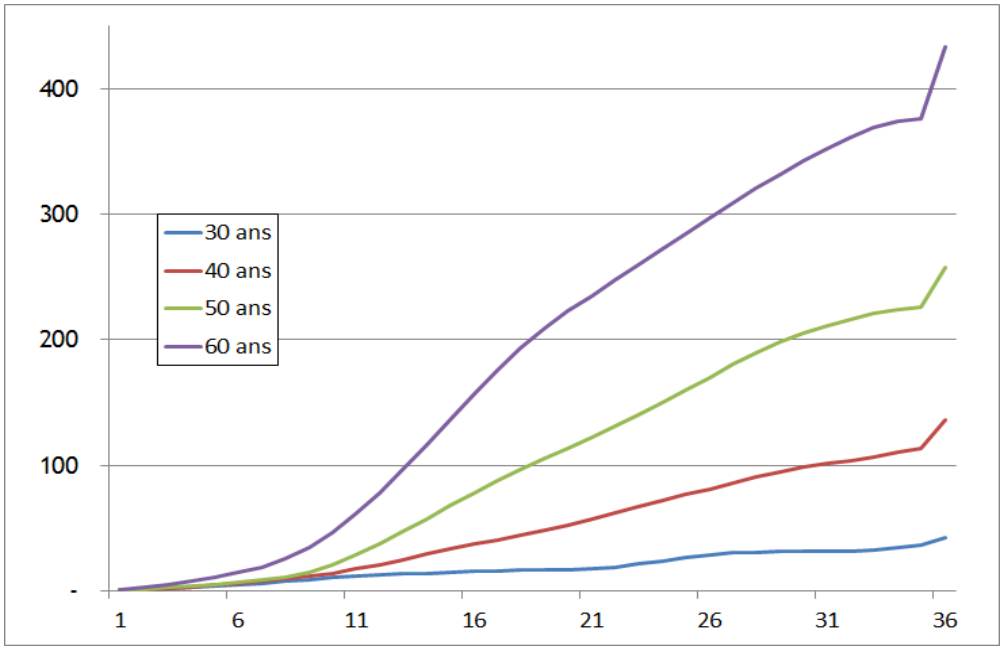
b. Validation du lissage

La validité du lissage est testée à travers un test du Khi-2.

En prenant un seuil de 95%, la région critique de notre étude est $\{W > 61\}$. La valeur de la statistique W sur notre table est de 55, ce qui nous conduit à accepter l'ajustement.

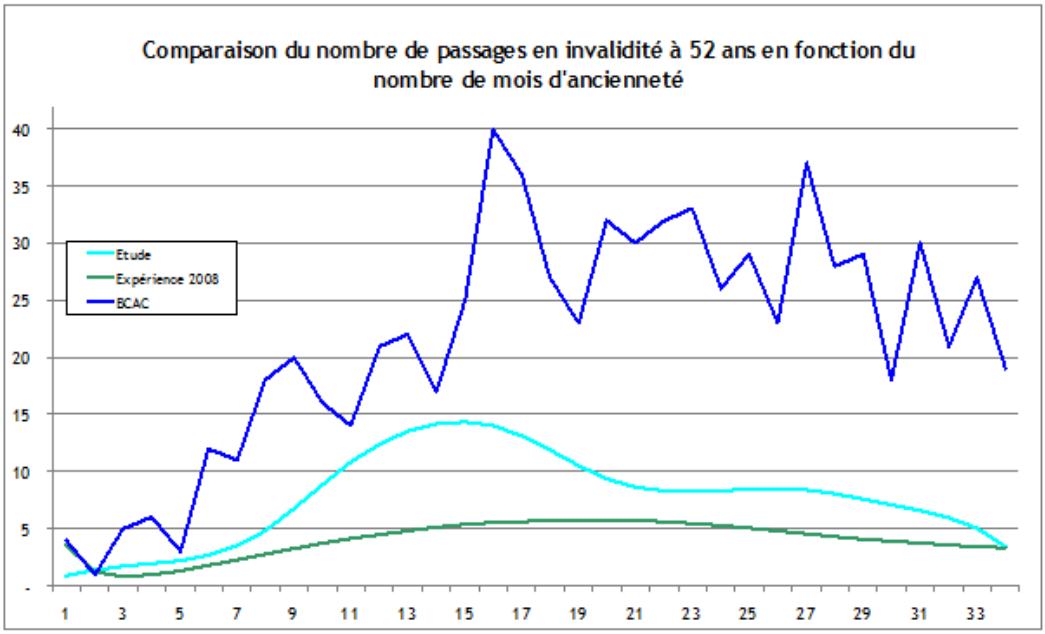
c. Résultat

Le graphique ci-dessous représente les nombres de passages en invalidité cumulés suivant l'ancienneté en mois d'incapacité à différents âges :



Nous constatons que plus l'incapable est âgé, plus il y a de passages en invalidité. Mise à part pour la courbe à 60 ans, les résultats sont très proches sur les 8 premiers mois pour les autres âges.

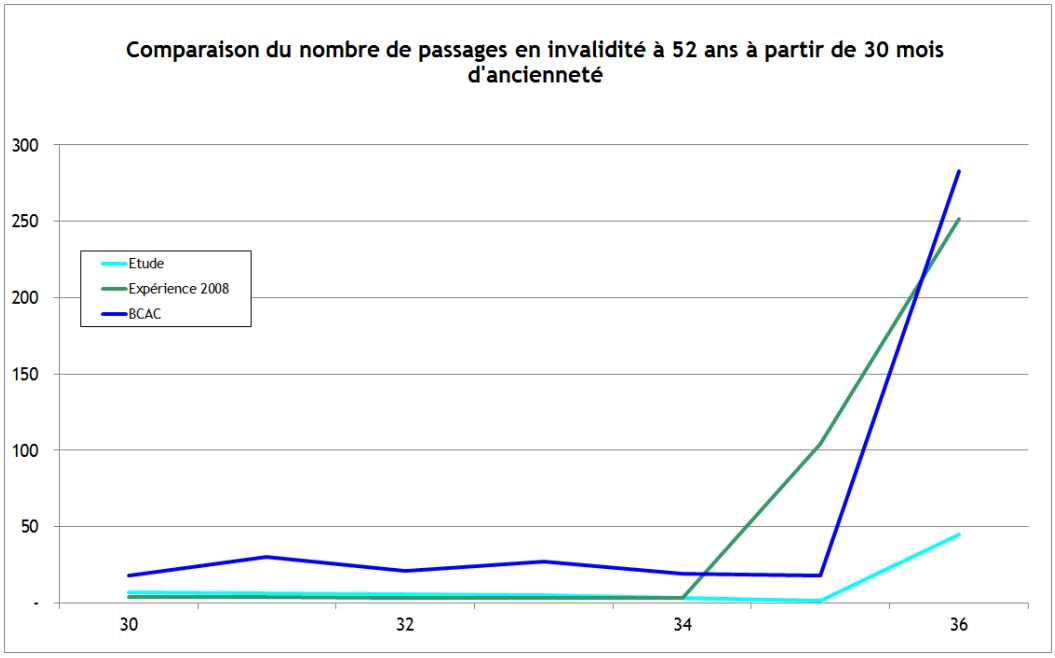
L'âge moyen d'entrée en invalidité de notre portefeuille étant de 52 ans, nous allons faire les comparaisons des différentes lois à cet âge :



Nous remarquons tout d'abord que la loi du BCAC n'est pas lissée et qu'elle est beaucoup plus élevée que notre courbe.

A partir de 2 mois d'ancienneté en incapacité, le nombre de passages en invalidité issu de l'expérience 2008 est inférieur au résultat de notre étude ; il y a moins de 5 passages d'écart jusqu'à 10 mois puis à partir du 19ème mois.

Regardons plus en détails les derniers mois d'ancienneté à 52 ans :



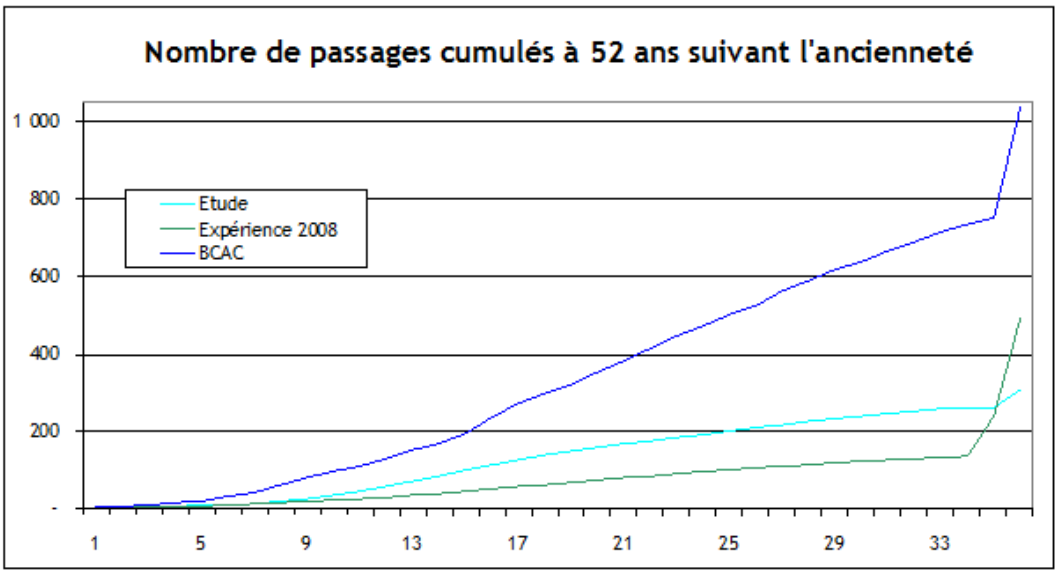
Nous remarquons qu'à 36 mois, le nombre de passages règlementaires et le nombre issu de l'expérience 2008 sont très proches et très élevés par rapport à notre nombre de transitions. En effet, nous avions définis par convention qu'à 36 mois et à tous âges, le taux de passage d'incapacité en invalidité était de 100%. Ainsi, pour un âge à la survenance d'incapacité de 52 ans et de 36 mois d'ancienneté, nous avons 45 incapables dans la table de maintien en incapacité et donc 45 transitions dans la table de passage d'incapacité en invalidité. Si nous observons ces 2 nombres (le nombre d'assurés âgés de 52 ans présents dans la table après 36 mois d'ancienneté) dans les autres tables, nous obtenons :

- pour la table règlementaire : 49 incapables et 283 passages
- pour la table d'expérience 2008 : 59 incapables et 251 passages

Notre convention parait plus cohérente, même si, quel que soit l'âge l'écart est important avec les autres tables.

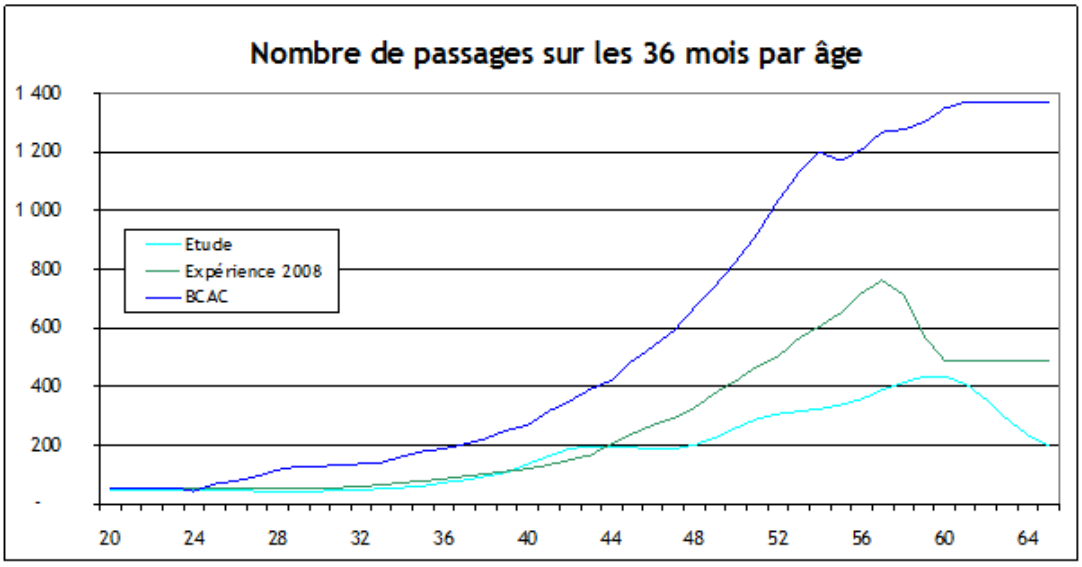
Sur les autres âges (les graphes sont en annexe 4a), nous remarquons que la courbe de la table d'expérience 2008 a toujours la même allure. En effet, cette table de transition a été construite sur les données de l'ensemble des sociétés participant à l'étude, tous âges confondus, puis, un ajustement a été effectué pour refléter le comportement du portefeuille Generali par âge.
En général, nous observons que notre courbe se situe entre les deux autres.

Le graphique suivant représente l'effectif cumulé des personnes passées en invalidité à 52 ans pour un effectif initial de 10 000 personnes en incapacité :



Nous observons que la table construite est la plus faible jusqu'à 10 mois d'ancienneté puis elle dépasse la courbe de 2008 tout en restant bien en dessous de celle du BCAC.
Sur les autres âges (les graphes sont en annexe 4b), nous remarquons que même si nos effectifs se situent généralement entre les deux autres courbes, les écarts sont plus ou moins importants suivant les âges et/ou les anciennetés.

Le graphique ci-dessous représente le nombre de passages en invalidité cumulés pour chaque âge à l'entrée :

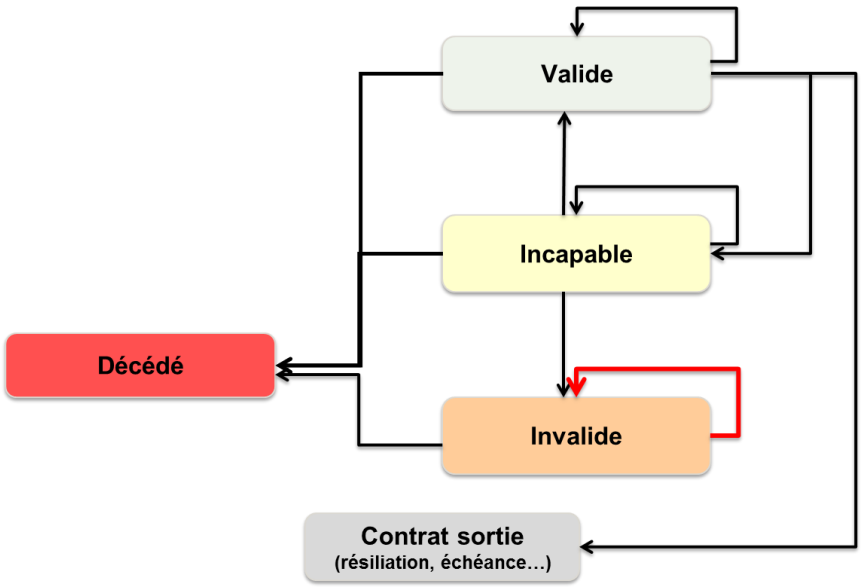


Nous constatons que la loi du BCAC est largement supérieure aux deux autres quel que soit l'âge à la survenance. Notre loi est, quant à elle, quasiment toujours inférieure à la loi d'expérience sauf pour la tranche d'âge de 40 à 43 ans. A partir de 60 ans, notre nombre de passages diminue, ce qui peut être expliqué par le faible nombre d'observations à ces âges.

Que ce soit pour le BCAC ou la table de 2008, à partir de 60 ans, le nombre de passage est cristallisé; c'est pourquoi nous observons une ligne droite à partir de cet âge. Cette méthode est appliquée pour le BCAC sur les moins de 23 ans et sur la table d'expérience pour les moins de 30 ans.

D. LA CONSTRUCTION DE LA TABLE DE MAINTIEN EN INVALIDITE

Pour notre modèle à état, il est nécessaire d'utiliser une table de maintien en invalidité pour calculer les provisions invalidité en Best Estimate et les cash-flows futurs, pour choquer le modèle sur le risque de maintien et pour définir la probabilité qu'un invalide reste invalide pour l'évolution de l'état des assurés :

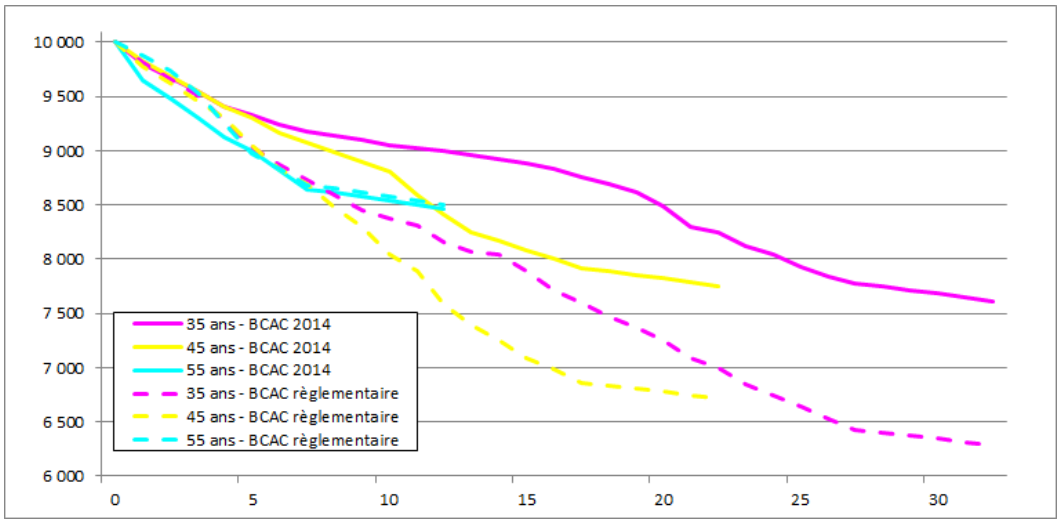


La table de maintien en invalidité se présente sous la forme d'un tableau à double entrée, avec en ligne l'âge à l'entrée en invalidité, et en colonne, l'ancienneté en année dans cet état. Dans chaque cellule, figure $l_{x,t}$ c'est à dire le nombre de personnes encore en invalidité sur les 10 000 de référence présentes au départ, pour l'âge à l'entrée x et une ancienneté de t années.

Comme vu précédemment, une base de 512 invalides ne comporte pas un volume de données suffisant pour construire une table de maintien fiable. Même en appliquant la méthode de glissement et en regroupement par ancienneté et âge, les résultats obtenus ne seraient pas robustes. Il ne nous est donc pas possible de construire une table de maintien en invalidité basée sur les données de notre portefeuille TNS.

Pour contourner ce problème et avoir une table qui se rapproche au plus de la sinistralité de notre portefeuille, nous choisissons d'utiliser la table de maintien en invalidité que le BCAC est en train de mettre à jour (table de mai 2014, dernière table réceptionnée à la date de cette étude) à l'aide de données de plusieurs assureurs dont celles de Generali. Cette table n'est ni définitive ni certifiée mais reflète certainement plus la réalité du risque que la table règlementaire construite en 1996.

Voici la comparaison des 2 tables de maintien en invalidité à différents âges :



Plus l'âge augmente et plus les écarts entre les deux tables diminuent, la nouvelle table majorant l'ancienne.

E. IMPACTS SUR LE CALCUL DES PROVISIONS

Pour estimer l'impact des nouvelles lois construites, nous recalculons les provisions incapacité/invalidité avec les tables règlementaires du BCAC, les tables d'expérience de 2008 et ces nouvelles tables.

1. Le calcul des provisions techniques

a. Notations

Pour les formules de calcul des provisions, nous posons les notations suivantes :

- x = âge à l'entrée en incapacité
- y = âge à l'entrée en invalidité
- anc = ancienneté en incapacité en mois à la date de calcul
- $\text{anc}_{\text{dernier règlement}}$ = ancienneté en incapacité en mois à la fin de période d'indemnisation du dernier règlement
- $\text{anc}_{\text{inval}}$ = ancienneté en invalidité en année à la date de calcul
- $l_{incap}(x,t)$ = effectif des personnes entrées en incapacité à l'âge x et toujours dans cet état au terme de t mois lu dans la table de maintien en incapacité
- $l_{inval}(y,t)$ = effectif des personnes entrées en invalidité à l'âge y et toujours dans cet état au terme de t années lu dans la table de maintien en invalidité
- $s(x,t)$ = effectif des personnes entrées en incapacité à l'âge x et qui deviennent invalides au cours du $t^{ème}$ mois lu dans la table de passage en invalidité
- f = durée de la franchise en nombre de mois
- i = taux d'intérêt technique
- g = taux de frais de gestion
- durée = nombre de mois d'incapacité restant
- montant garantie = montant de la prestation annualisée (pour les rentes invalidité ou les indemnités journalières)
- CI = montant du capital invalidité souscrit
- Pi = portion des invalidités de 1ère catégorie (donnée théorique)
- Montant garantie = montant de la prestation annuelle définit par l'assurée à la souscription

b. Le calcul des PSAP

Les PSAP correspondent aux Provisions pour Sinistres A Payer : les provisions des sinistres incapacité en cours de service. Elles se calculent en fonction de l'âge à l'entrée en incapacité et de l'ancienneté sur les sinistres ouverts avec un règlement depuis moins de 6 mois (dépassé ce délais, nous considérons ce sinistre comme clos). Le coefficient est calculé de la façon suivante :

$$C_{incap}(x,t,f) = \begin{cases} \dfrac{1}{(1+i)^{f-t}} \displaystyle\sum_{k=1}^{durée-f} \dfrac{1}{(1+i)^k} * \dfrac{l_{incap}(x, f+k)}{l_{incap}(x, f)}, \text{ pour} t < f \\ \displaystyle\sum_{k=1}^{durée-t} \dfrac{1}{(1+i)^k} * \dfrac{l_{incap}(x,t+k)}{l_{incap}(x,t)}, \text{ pour} t \geq f \end{cases}$$

La provision d'incapacité est obtenue par

PSAP = C_{incap} (age, anc, f) * Montant garantie * (1 + g)

c. Le calcul de l'invalidité en attente

En cas d'invalidité, nos produits peuvent garantir soit un versement de rentes soit le paiement d'un capital. Lorsque l'assuré est en incapacité et qu'il a souscrit une des deux garanties invalidités, en plus de la PSAP, nous calculons :

- soit une PSAPR : Provisions pour Sinistres à Payer pour Rente d'invalidité en attente (nous l'appellerons PSAPR dans la suite)
- soit un CIPR : Capital Invalidité Probable (nous l'appellerons CIPR dans la suite)

i. PSAPR

Nous calculons une PSAPR pour les incapables ayant souscrits la garantie rente invalidité. Cette provision prend en compte la probabilité que nous ayons à régler une rente à l'assuré s'il passe en invalidité.

Le coefficient de provisionnement est déterminé par :

$$C_{RA}(x,t) = \sum_{k=1}^{11} \frac{s(x,t+k)/l_{incap}(x,t)}{(1+i)^k} * C_{inval}(x+1,0)$$
$$+ \sum_{k=12}^{23} \frac{s(x,t+k)/l_{incap}(x,t)}{(1+i)^k} * C_{inval}(x+2,0)$$
$$+ \sum_{k=24}^{durée-t} \frac{s(x,t+k)/l_{incap}(x,t)}{(1+i)^k} * C_{inval}(x+3,0)$$

La provision de rente en attente est alors obtenue de la façon suivante (le C_{inval} étant définit plus bas) :

PSAPR = $C_{RA}(age, anc)$ * Montant garantie * (1 + g)

A noter que cette provision utilise la table de maintien en incapacité, la table de passage en invalidité et la table de maintien en incapacité.

ii. CIPR

Nous calculons un CIPR pour les assurés incapables ayant souscrits la garantie capital invalidité. Cette provision prend en compte la probabilité que nous ayons à régler le capital à l'assuré s'il passe en invalidité.

Le coefficient de provisionnement est déterminé par :

$$C_{CI}(x,t) = \frac{s(x,t)}{l_{incap}(x,t)} * (1 - Pi)$$

Le montant de la provision est obtenu comme suit :

$$CIPR = C_{CI}(age, anc) * CI * (1+g)$$

d. Le calcul des PMR

Les PMR correspondent aux Provisions Mathématiques pour Rentes en invalidité (nous les appellerons PMR dans la suite). Elles correspondent aux provisions des sinistres invalidité en cours de service, en fonction de l'âge d'entrée en invalidité et de la durée courue (en année) dans cet état. C'est le montant qui va servir à couvrir les rentes versées après que l'assuré ait été déclaré invalide que ce soit après 3 ans d'incapacité ou immédiatement.

Le coefficient de provisionnement pour un invalide est de :

$$C_{inval}(y,t) = \sum_{k=1}^{age\max - y - t} \frac{1}{(1+i)^k} * \frac{l_{inval}(y,t+k)}{l_{inval}(y,t)}$$

La provision d'invalidité est alors obtenus par :

$$PMR = C_{inval}(age, anc_{inv}) * \text{ Montant garantie } * (1+g)$$

e. Le coefficient d'abattement pour l'incapacité

Pour prendre en compte la probabilité que l'assuré soit toujours en incapacité entre la date de fin de période d'indemnisation (qui correspond au dernier règlement) et la date de calcul de la provision, nous appliquons un coefficient d'abattement. Ce coefficient (compris entre 0 et 1) permet de prendre en compte :

- Soit une reprise de travail de l'assuré ou un passage en invalidité sans que l'assureur en ait déjà connaissance
- Soit un retard en gestion

A noter que si la date du dernier règlement correspond à la date de calcul alors ce coefficient sera égal à 1.

Ce coefficient est calculé de la manière suivante :

$$\text{Coeff abattement}(x, anc) = \frac{l_{incap}(x, anc) + \sum_{k=(\text{dernier règlement})+1}^{\text{date calcul}} s(x, k)}{l_{incap}(x, anc_{\text{dernier règlement}})}$$

Ce coefficient est appliqué sur les PSAP, PSAPR et CIPR.

Nous obtenons les formules suivantes pour ces provisions :

PSAP finale = PSAP * Coefficient d'abattement

PSAPR finale = PSAPR * Coefficient d'abattement

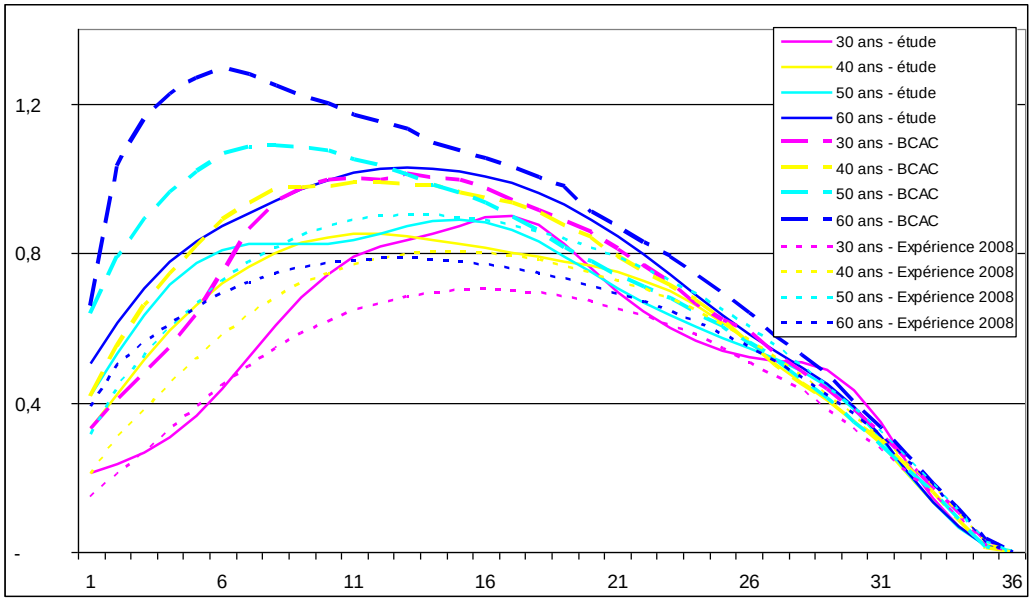
CIPR final = CIPR * Coefficient d'abattement

Nous parlons alors de provisions affinées.

2. La comparaison des montants de provisions

Regardons en détails, les impacts des nouvelles tables pour 1€ de prestation à différents âges pour les provisions les plus importantes.

a. L'impact sur les PSAP



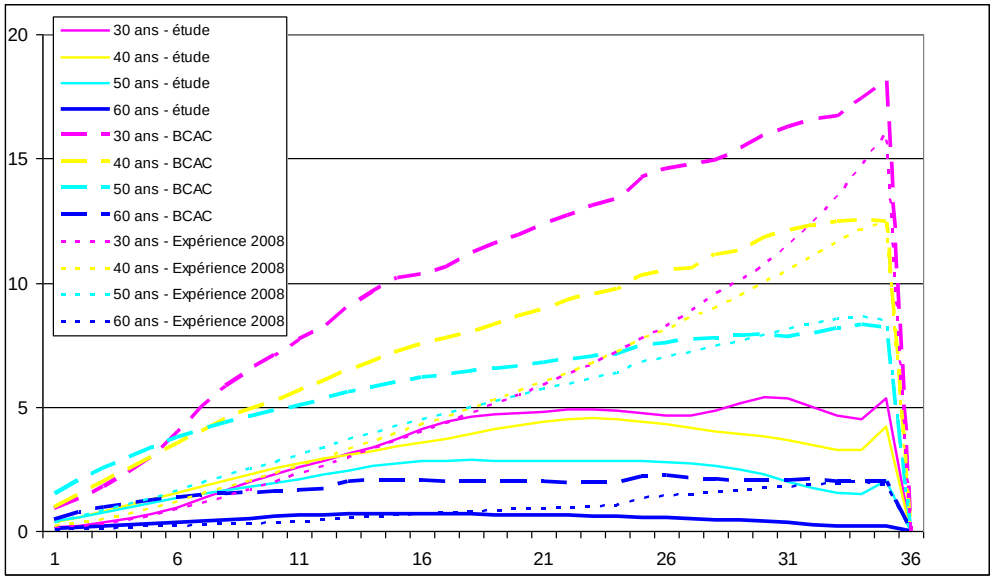
Ce graphique représente le montant de la PSAP pour le versement d'une IJ à 1€ pour un assuré incapable en fonction du nombre de mois d'ancienneté. Une comparaison des provisions est effectuée pour différents âges à la survenance et entre les différentes lois (nos lois lissées en

courbes continues, les lois du BCAC en courbes discontinues et les lois d'expérience de 2008 en courbes pointillées).
Nous observons tout d'abord que pour un âge équivalent, les PSAP du BCAC sont toujours plus élevées de 20% en moyenne que les PSAP de notre étude et celles-ci augmentent d'environ 4% par rapport aux provisions calculées avec la table d'expérience de 2008. Les écarts les plus importants se situent sur les anciennetés inférieures à un an.
Plus l'ancienneté augmente, plus l'écart entre les tables diminue. A partir de 28 mois d'ancienneté, l'ensemble des courbes se rejoignent pour arriver à un montant de 0€ à 36 mois. En effet, à partir de 36 mois d'incapacité, c'est la provision mathématique de rente en invalidité qui prend le relais car, nous supposons qu'au bout de 36 mois maximum d'incapacité, le sinistré passe en invalidité.

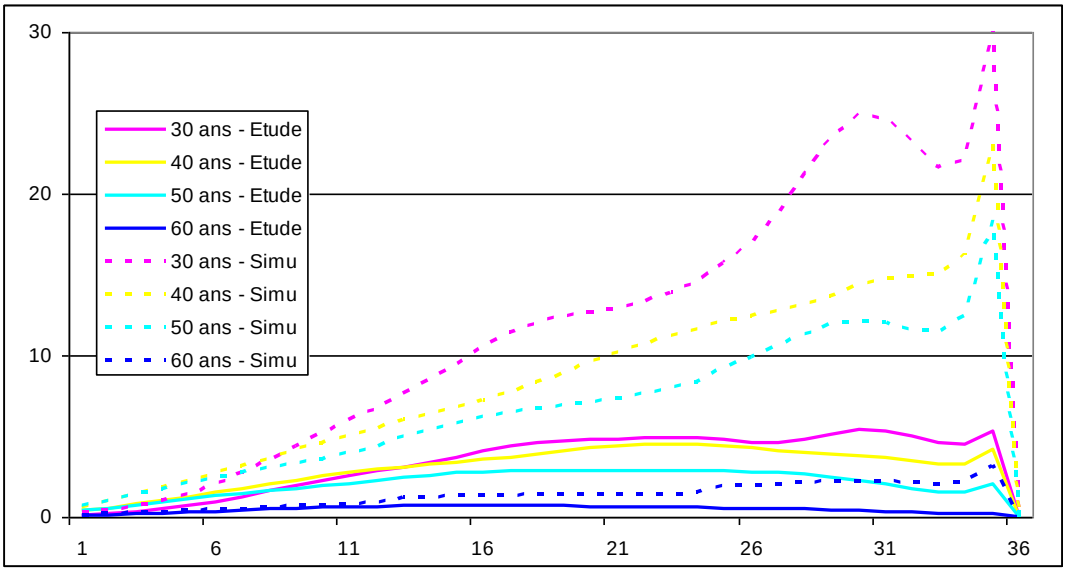
En générale, la provision augmente avec l'âge quel que soit la loi utilisée. Cependant, nous remarquons, par exemple, que pour certaines anciennetés, la PSAP à 30 ans est plus importante que celle à 40 ans aussi bien avec la loi du BCAC qu'avec notre loi. Cette constatation peut s'expliquer par le fait que nous avons choisi de lisser la loi et non les provisions comme cela est fait dans certaines études.

b. L'impact sur les PSAPR

Cette provision est difficile à analyser car elle combine les 3 tables de morbidité pour son calcul.
Le graphique ci-dessous représente le montant de la PSAPR par nombre de mois d'ancienneté et pour différents âges pour une prestation d'1€. Une comparaison est effectuée entre les provisions calculées avec nos lois lissées (courbes continues), avec les lois règlementaires du BCAC (courbes discontinues) et les lois d'expérience de 2008 (courbes en pointillées) :



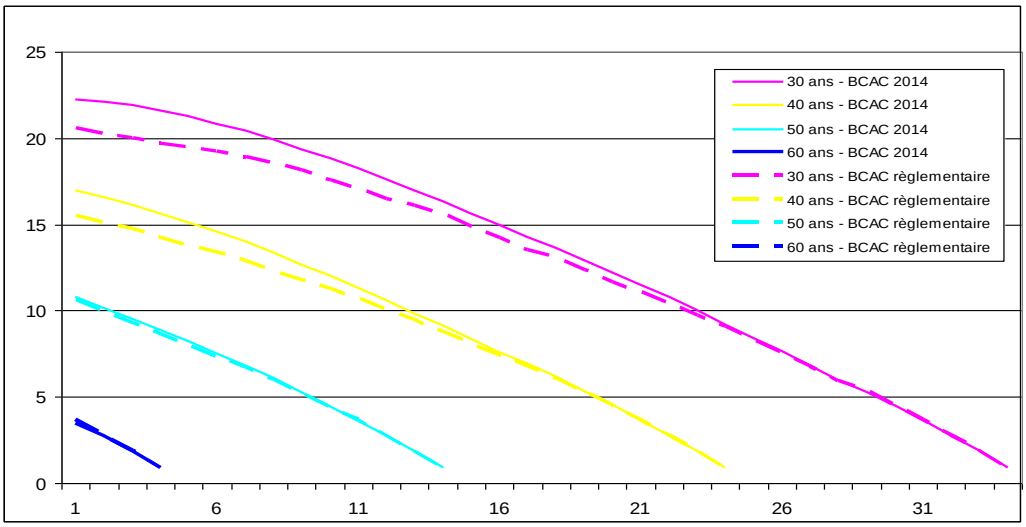
De façon générale, nous remarquons que plus le sinistré est jeune, plus sa provision pour rente en attente est importante. En effet, plus un assuré est jeune, plus sa durée possible d'invalidité est longue et donc plus sa provision de passage est importante.
Nous observons que quel que soit l'âge et quel que soit l'ancienneté, les tables du BCAC engendrent les montants de provisions les plus importants, ils sont notamment 3 fois plus élevés que les PSAPR de notre étude.
Les PSAPR calculées avec notre étude sont les plus faibles et diminuent quasiment de moitié par rapport aux provisions calculées avec les tables 2008.



Ce graphique représente toujours les PSAPR pour 1 € de prestation calculées avec les lois de notre étude (en courbes continues) et le montant des provisions calculées avec la table de passage issue de notre étude où nous avons remplacé le 36ème mois par celui du BCAC (calculs que nous appelons simulations représentés par les courbes discontinues).
Nous observons ainsi que les PSAPR simulées sont bien croissantes avec l'âge mais surtout qu'elles sont beaucoup plus importantes que celles de notre étude. Le nombre de passages au 36ème mois est donc très important car il a un impact très significatif sur le montant de la provision calculé à chaque âge et à chaque ancienneté. Ceci s'explique par la formule du coefficient de rente en attente qui somme tous les termes de la table entre l'ancienneté et le 36ème mois. Les calculs effectués avec nos tables d'expérience sont les plus faibles et s'expliquent par la convention que nous avons adoptée pour définir la valeur du nombre de passage au 36ème mois d'ancienneté.

c. L'impact sur les PMR

Le graphique suivant représente le montant des PMR pour une rente de 1€ à différents âges à la survenance selon l'ancienneté (en nombre d'années). Une comparaison est effectuée entre les provisions calculées avec la loi non définitive 2014 du BCAC (courbes continues) et avec la loi règlementaire du BCAC (courbes en pointillées) :



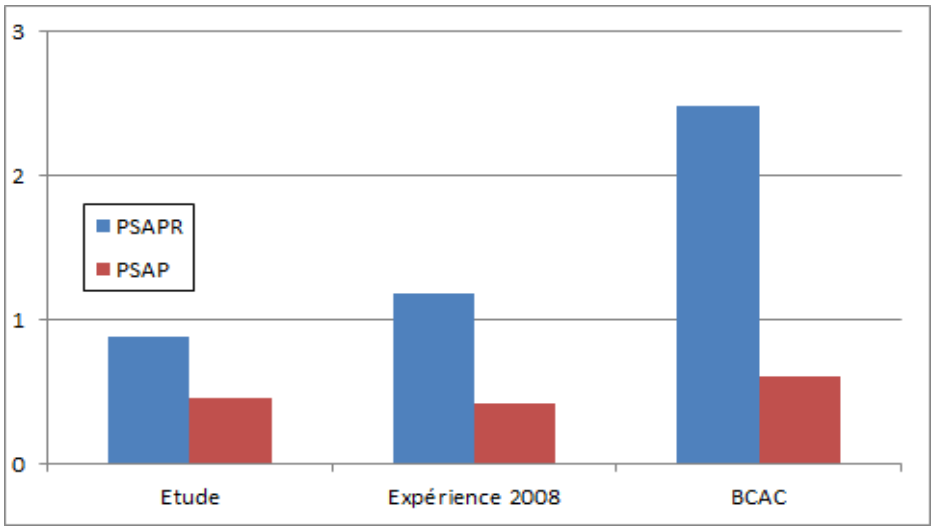
L'invalidité prenant fin, en général, aux 65 ans de l'assuré, plus le sinistré invalide est jeune plus sa durée probable d'invalidité peut être longue, plus sa provision va être importante. C'est pourquoi le montant de la provision diminue lorsque l'âge augmente.

Pour les 50 et 60 ans, le montant des provisions est très proche entre les 2 tables. Sur les autres âges, la loi actualisée majore légèrement le montant des PMR d'environ 5%.

d. L'impact réel sur notre portefeuille

A présent, nous allons évaluer l'impact réel du changement de tables sur les provisions de notre portefeuille. Nous récupérons, ainsi, nos assurés sinistrés en fin d'année et nous calculons têtes par têtes les provisions affinées avec les tables règlementaires du BCAC (calcul n°1), avec les tables d'expériences 2008 (calcul n°2) et avec les tables issues de notre étude (calcul n°3).

Le graphique ci-dessous compare le nombre moyen d'années provisionnées (Montant des provisions / Montant des prestations annualisées) sur notre portefeuille pour les sinistres incapacité en cours :



Avec les tables d'expériences (celle de 2008 ou la nôtre), nous provisionnons en PSAP environ 5 mois d'incapacité et plus de 7 mois avec la table règlementaire du BCAC. L'ancienneté moyenne en incapacité étant de 6 mois, nous estimons qu'un incapable va rester en moyenne 11 mois dans cet état avec les tables d'expérience et 13 mois avec la table du BCAC.

Concernant la PSAPR, nous provisionnons moins d'un an avec notre table, un peu plus d'un an avec la table d'expérience de 2008 et deux ans et demi avec la table règlementaire du BCAC.

Concernant les rentes d'invalidité en cours, nous provisionnons 7,5 ans en moyenne avec la table du BCAC règlementaire et presque 8 ans avec la table du BCAC 2014; l'âge moyen actuel du portefeuille des invalides étant de 55 ans, nous estimons qu'en moyenne un invalide reste dans cet état jusqu'à 63 ans donc jusqu'à son départ en retraite.

A présent, nous comparons les PSAP (pour l'incapacité des incapables en cours), les PSAPR (pour les rentes d'invalidité en attente pour les incapables en cours), les CIPR (pour les capitaux invalidité probables pour les incapables en cours) et les PMR (pour les invalidités en cours) de notre portefeuille suivant les différentes tables:

	Calcul n°1	Calcul n°2	Calcul n°3
Maintien en incapacité	BCAC règlementaire	Expérience 2008	ETUDE
Passage en invalidité	BCAC règlementaire	Expérience 2008	ETUDE
Maintien en invalidité	BCAC règlementaire	BCAC règlementaire	BCAC 2014
PSAP (k €)	57 551	39 094	43 550
PSAPR (k €)	146 079	68 725	51 438
CIPR (k €)	118	90	30
PMR (k €)	120 615	120 615	125 871
TOTAL (k €)	324 363	228 523	220 889

Le calcul n°1 correspond aux calculs des provisions incapacité et invalidité règlementaires.
Le calcul n°2 prend en compte les tables d'expérience 2008 pour les provisions des incapables et la table règlementaire pour les invalides.
Le calcul n°3 utilise pour les incapables les tables de notre étude et pour les invalides la table BCAC 2014 non définitive.

Le calcul n°3 diminue le montant des provisions de plus de 30% par rapport aux tables réglementaires du BCAC avec un écart très important sur les PSAPR et une légère augmentation sur les PMR. La PSAPR est la provision la plus sensible car, c'est elle, qui a le plus gros impact sur le résultat final.

Entre le calcul n°3 et n°2, l'écart au global est faible atteignant les 3% ; le calcul n°3 minorant les provisions. Le risque invalidité se dégrade légèrement avec une hausse des PMR de 4% alors que les PSAP augmentent de 10% et que les PSAPR diminuent fortement de 33%.

L'application du coefficient d'abattement dans le calcul diminue au global d'environ 10% le montant des provisions issues de notre étude. Il n'y a pas d'impact sur les PMR par contre l'abattement est d'un peu plus de 20% sur les PSAP et PSAPR.

En regardant en détail le montant des provisions entre le calcul n°2 et n°3, nous observons l'impact du changement de table :

- touche toutes les anciennetés pour l'invalidité en augmentant légèrement les provisions sur toutes les survenances
- se concentre surtout sur la dernière année d'incapacité appuyée par un impact de la PSAPR très important, même si, sur la première année d'incapacité, notre étude majore de 7% les provisions issues des tables 2008.

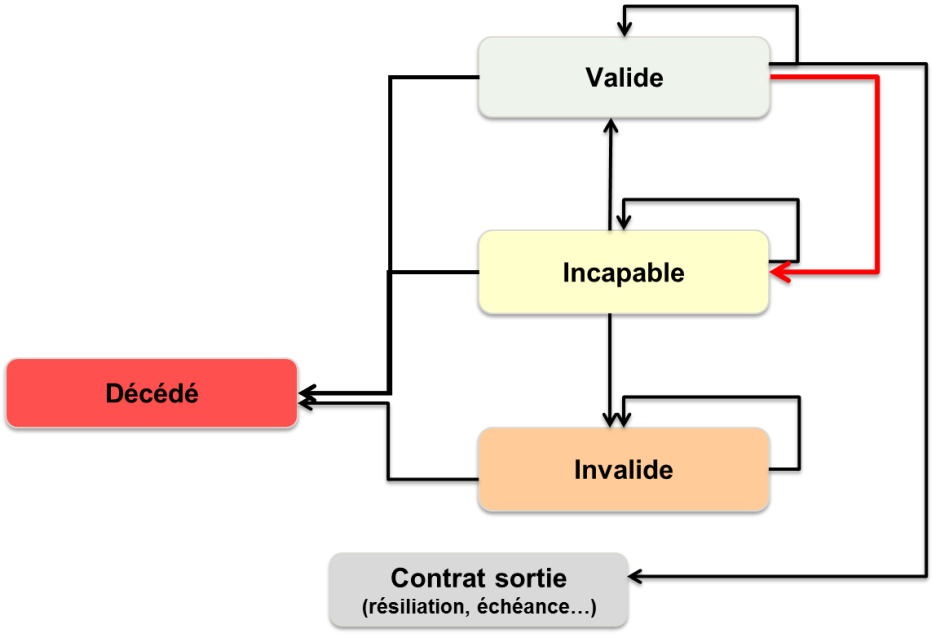
Ces différentes comparaisons sont intéressantes car elles montrent que :

- la sinistralité propre à notre portefeuille diverge largement de celle établit par le BCAC (sur une population de salariés)
- un ajustement était nécessaire par rapport aux tables construites en 2008 pour calculer des provisions en Best Estimate et pour une modélisation plus fine

Cela justifie donc notre choix de vouloir construire des tables d'expérience sur notre propre portefeuille d'assurés.

IV. CONSTRUCTION DES LOIS D'ENTREE EN INCAPACITE

Comme vu précédemment, dans le cadre du modèle interne, nous construisons les hypothèses pour le modèle à états. Chaque état et chaque passage à un état est modélisé. Pour modéliser le passage de l'état valide à l'état incapable nous avons donc besoin d'une fréquence d'entrée en incapacité (aussi appelée loi d'incidence) par âge: c'est le but de ce chapitre.
Cette loi nous servira également pour calibrer les chocs à appliquer au modèle sur le risque d'incidence.



A noter que nous n'utiliserons pas tout à fait les mêmes données que celles utilisées pour la construction des lois précédentes. En effet, nous avons besoin de la base des contrats ayant souscrits l'incapacité et des sinistres associés. Or, dans l'extraction des contrats de l'ancien système de gestion, nous n'avons pas la distinction entre l'incapacité et l'invalidité souscrites (sachant qu'un assuré peut souscrire l'incapacité sans souscrire l'invalidité). Nous n'utilisons donc que les données présentes dans le système actuel c'est-à-dire à partir de 2008.

A. LES DONNEES

1. La présentation des fichiers sources

Nous utilisons les tables SAS, issues de l'applicatif, extraites à fin octobre 2013 :

- La table des Personnes pour récupérer la date de naissance de l'assuré

- Les tables des Contrats/Garanties pour ne sélectionner que les contrats ayant souscrits une garantie incapacité et pour constituer une vision des contrats en stock à chaque année d'observation (à l'aide de la date d'effet du contrat et de la date de sortie du contrat)
- La table des Sinistres qui contient l'ensemble des règlements effectués par garantie et contrat

La clé de jointure entre toutes ces tables est le numéro de contrat.

2. Les traitements des données

Le traitement des données s'effectue sous SAS sur l'ensemble des fichiers sources.

a. Préparation des bases

Nous effectuons une préparation des bases pour ne conserver que les enregistrements nécessaires à notre étude :

- Pour l'ensemble des tables, nous ne conservons que les lignes relatives à l'incapacité et aux produits PPL/PGM et ATOLL TNS
- Pour la table contrat, nous supprimons les contrats sans effet (ces contrats ne doivent pas être pris en compte dans nos effectifs)
- Pour la table sinistre nous supprimons les sinistres sans règlement et les règlements annulés (pour ne prendre en compte que des sinistres où il y a eu effectivement un règlement et ainsi ne conserver que les sinistres dépassants leur période de franchise)

b. Traitement de la base des sinistres

Pour la base des sinistres, nous devons faire les traitements suivants :

- Sélectionner une seule ligne par numéro de contrat, numéro de sinistre, garantie, date de survenance, dates minimum et maximum de la période de règlement du sinistre
- Calculer la durée du sinistre :
    - Date de début du sinistre = Date de survenance du sinistre
    - Date de fin du sinistre = Date de fin de période de règlement maximum du sinistre
- Calculer l'année de survenance du sinistre

La base des sinistres contient ainsi une ligne par couple contrat/sinistre.

c. Construction des tables de bases

Nous créons deux bases pour notre étude :

- La base des sinistres contenant l'ensemble des sinistres incapacité par contrat, produit, année de survenance, date de début et de fin du sinistre, âge à la survenance

- La base contrat/garantie contenant l'ensemble des contrats ayant souscrit une garantie incapacité, le produit, la date d'effet du contrat, la date de sortie du contrat, la date de naissance de l'assuré

d. Retraitements des données

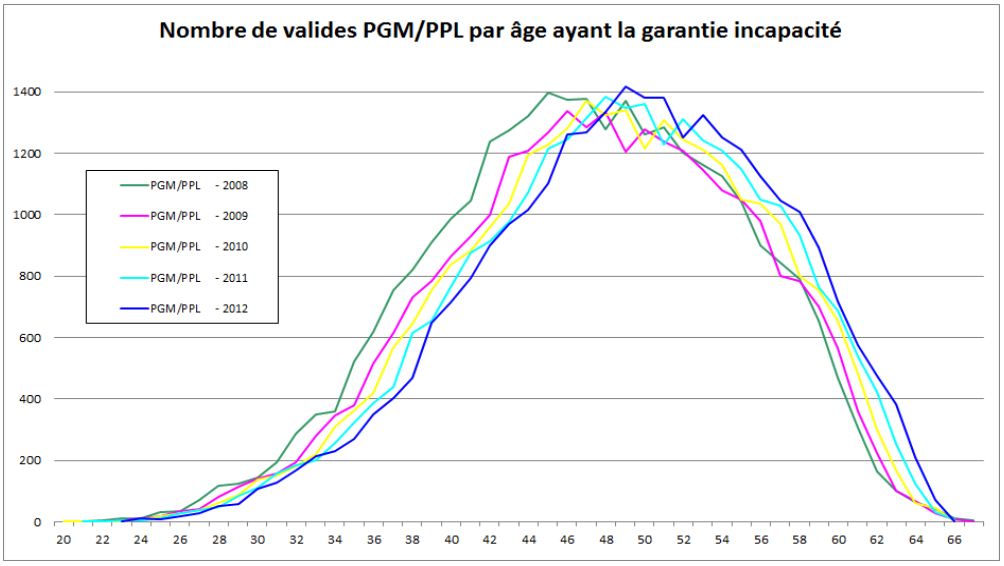
Dans le cadre de la QDD, nous faisons quelques contrôles de cohérence, dont la vérification qu'un assuré sinistré est bien dans la base des expositions aux risques (donc qu'il a bien la garantie incapacité souscrite). Ce contrôle a mis en évidence que ce n'était pas toujours le cas. Après vérification dans le système de gestion, il se trouve que ces contrats avaient bien souscrits une garantie incapacité à l'origine, qu'ils ont eu un sinistre incapacité, puis il y a eu une modification des garanties souscrites pour retirer la garantie incapacité. En effet, la table SAS utilisée donne une vision des garanties à l'instant t de l'extraction, sans l'historique des mouvements de garanties.
Pour ne pas biaiser l'étude et pour conserver quand même ces sinistres, nous décidons de rajouter ces contrats sinistrés à la base des contrats ayant souscrits la garantie incapacité.

Pour obtenir les statistiques annuelles, nous devons reconstituer une base contrat/garantie et une base sinistre à chaque date de calcul (01/01/N). Ainsi, pour chaque contrat, à l'aide de la date d'effet du contrat et de sa date de sortie, nous pouvons reconstituer son état (contrat en cours ou contrat sorti ou affaire nouvelle) et l'âge de l'assuré à chaque date de calcul. De même pour les sinistres, à l'aide de la date de survenance, nous pouvons identifier les sinistres entrants en incapacité et à l'aide des dates de début et dates de fin de sinistre, nous pouvons reconstituer les stocks de sinistres à chaque date de calcul et l'âge des sinistrés à la survenance.

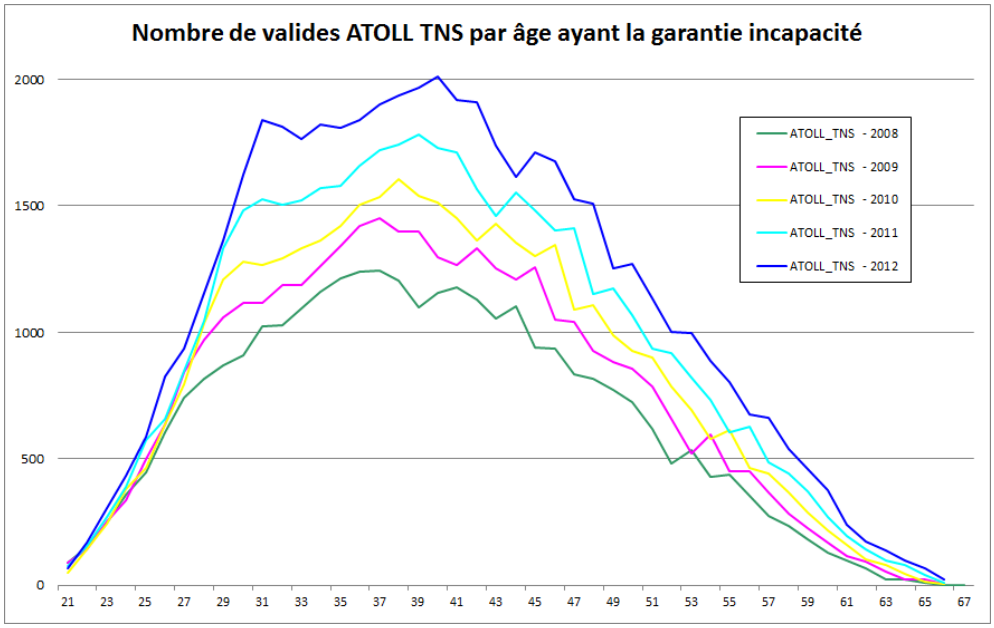
3. Les statistiques sur le portefeuille obtenu

a. Nombre de contrats en cours

Voici les contrats en cours par produits pour chaque date d'observation (de 2008 à 2012, nous avons supprimé l'année 2013 des statistiques car l'année n'etait pas entière) :



Nous remarquons que chaque année la courbe glisse vers la droite ce qui représente le vieillissement à chaque année de la population assuré. La croissance est plutôt stable pour PGM/PPL.



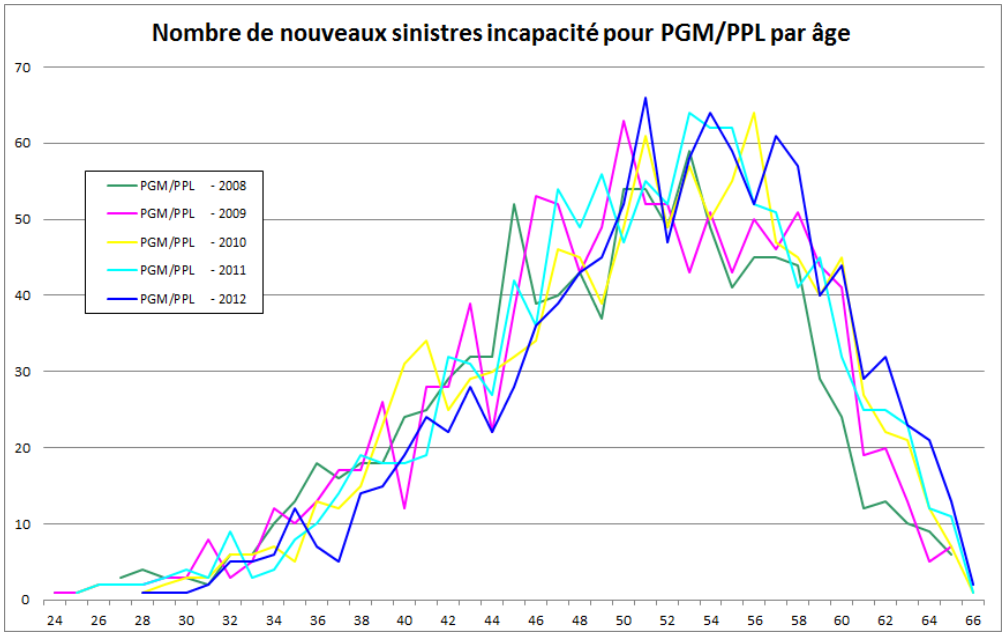
Nous remarquons qu'ATOLL TNS est en fort développement d'année en année, la croissance est de +70% entre 2008 et 2012. Les âges des assurés restants plutôt stables.

La population d'ATOLL est plus jeune que le portefeuille de PGM/PPL : l'âge moyen est de 40 ans pour ATOLL et de 48 ans pour PGM/PPL.

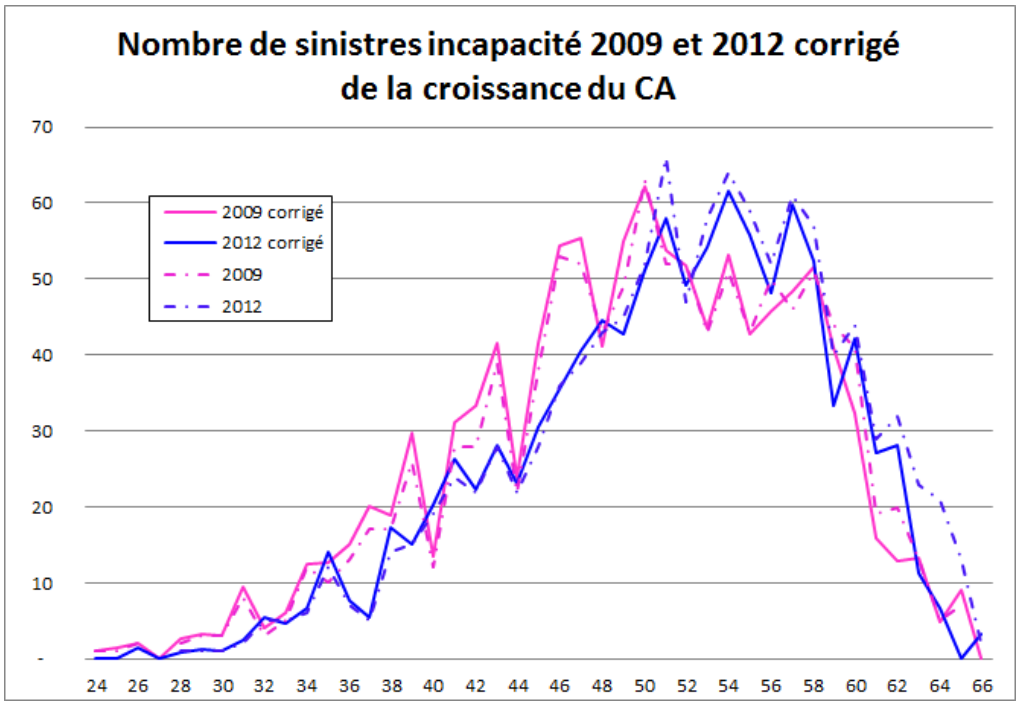
Quel que soit le produit, pour les âges extrêmes nous avons moins de 200 contrats observés. Il sera donc délicat d'obtenir des taux fiables pour ces âges extrêmes. Pour pallier à cette problématique, nous allons donc construire une loi par classe d'âges.

b. Nombre de nouveaux sinistres

Le graphique ci-dessous représente le nombre de nouveaux sinistres incapacité PGM/PPL par âge à la survenance sur les différentes années d'observation :

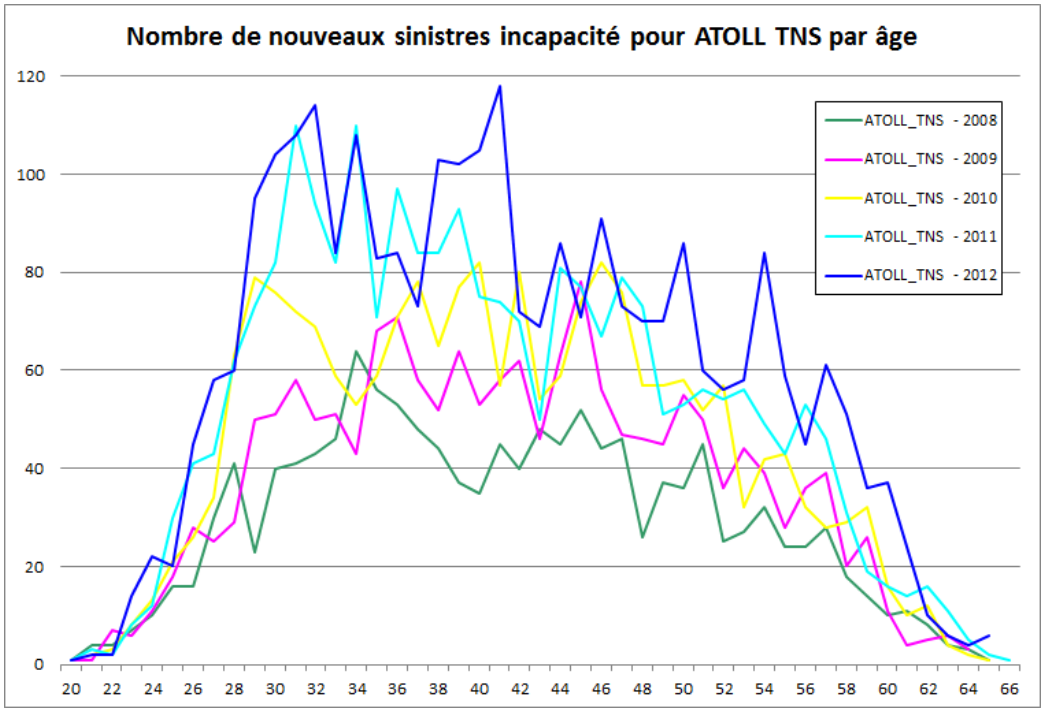


Nous observons une certaine stabilité de la sinistralité sur PGM/PPL au fils des années (+8% entre 2008 et 2012).

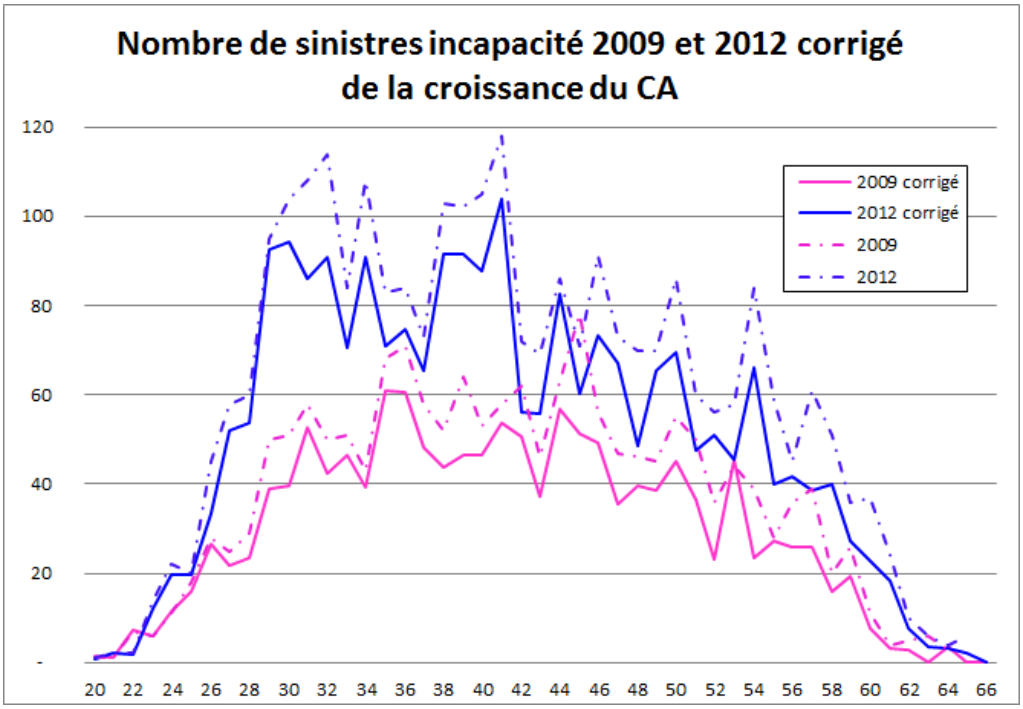


Même en corrigeant le nombre de nouveaux sinistres de la croissance du chiffre d'affaires, la variation de sinistralité entre 2009 et 2012 n'augmente pas sauf sur certains âges (de 53 ans à 59 ans puis de 61 à 63 ans).

De même, voici le graphique du nombre de nouveaux sinistres incapacité d'ATOLL TNS :



Sur ATOLL, nous observons une augmentation importante du nombre de sinistres suivant les années d'observation (+106% entre 2008 et 2012).



Après avoir corrigé le nombre de nouveaux sinistres de la croissance du chiffre d'affaires, nous constatons toujours une dérive de la sinistralité sur ATOLL quel que soit l'âge du sinistré entre 2009 et 2012.

L'âge moyen des sinistrés entrants sur ATOLL est de 41 ans contre 50 ans pour PGM/PPL.

La population des deux portefeuilles est donc bien différente car leur sinistralité diverge aussi bien sur les âges moyens que sur la fréquence d'apparition des sinistres qui ne cesse d'augmenter sur ATOLL TNS. Il est donc nécessaire de construire des lois d'incidence par produit.

Au vu du nombre d'observation par âge, nous décidons de faire notre étude sur des tranches d'âges pour plus de robustesse, puis une fois la courbe lissée, nous en déduirons des taux par âge.

Etant donné le nombre de sinistres observés, nous ne pourrons faire aucune segmentation autre que l'âge et le produit dans notre étude.

B. LA METHODE

Nous utilisons des données propres à chaque produit afin de tenir compte de la spécificité de chacun.

1. Les taux bruts

Nous souhaitons un taux de fréquence d'entrée en indemnisation, donc une fois que la franchise est dépassée. Ce taux se définit comme la probabilité pour un assuré exposé au risque pendant l'année d'entrer en indemnisation d'incapacité au moins une fois pour un sinistre survenant cette même année. Si l'assuré a plusieurs sinistres dans l'année, nous le comptons qu'une seule fois. Nous devons donc construire deux bases : une base des personnes exposées au risque c'est-à-dire des valides ayant souscrits la garantie et une base des nouveaux sinistres de l'année.

La fréquence empirique sur une année d'observation N pour un âge x ou une classe d'âge x est la suivante :

$$\text{Fréquence entrée inca}\mathbf{p_x^N} = \frac{\text{Nombre de nouveaux sinistres}_\mathbf{x} \text{ sur } \mathbf{N}}{\text{Nombre d'assurés valides}_\mathbf{x} \text{ au } 01/01/\mathbf{N}} = \frac{\mathbf{S_x^N}}{\mathbf{V_x^N}}$$

Durant toute l'année N, un contrat peut avoir plusieurs états différents :

- Soit il est en cours au 01/01/N, avec une garantie incapacité souscrite; il est donc exposé au risque d'entrée en incapacité sur l'année N : A_x^N . Le contrat est en cours l'année N si $\text{date d'effet du contrat} \leq 31/12/\text{N-1}$ et, soit la date de sortie du contrat n'est pas renseignée, soit $\text{date de sortie du contrat} > 31/12/\text{N-}\mathbf{1}$
- Soit c'est une affaire nouvelle sur N avec une garantie incapacité souscrite et il sera exposé au risque tout le reste de l'année N : B_x^N . Le contrat est une affaire nouvelle en N si l'année d'effet du contrat est N

- Soit il est sinistré au 01/01/N et nous faisons l'hypothèse qu'il le reste toute l'année (il ne fait donc pas partie des assurés valides au 01/01/N) : C_x^N . Ce sinistre est considéré en stock pendant l'année N si $\text{date de début du sinistre} \leq 31/12/\text{N-1} \leq \text{date de fin du sinistre}$

- Soit il entre en indemnisation durant l'année N et il était déjà sinistré au 01/01/N (il ne fait donc pas partie des assurés valides au 01/01/N) : E_x^N . Dans ce cas, comme le contrat est déjà sinistré au 01/01/N, cet assuré ne fait pas parti de la base des valides et nous le supprimons donc de la base des sinistres entrants

- Soit il entre en indemnisation durant l'année N : D_x^N . A noter qu'un contrat ayant plusieurs sinistres la même année ne sera compter qu'une seule fois. Le sinistre entre en indemnisation en N si l'année de survenance est N

- Soit il entre en indemnisation durant l'année N, il est non sinistré au 01/01/N et il n'est pas dans l'effectif des contrats exposés au risque : F_x^N . Ce cas apparaît lorsque qu'un contrat souscrit la garantie incapacité, puis un sinistre survient, puis la garantie incapacité n'est plus garantie au contrat. Etant donné que la table garantie utilisée est une vision des garanties à un instant t nous ne pouvons pas reconstituer l'ensemble des garanties à un instant autre que t. Ainsi pour ne pas compter un sinistre qui n'existe pas dans la base des valides, nous décidons de rajouter à la base des valides tous les sinistres dans ce cas

- Soit le contrat sort : cet état n'est pas modélisé

Notre base en N des nouveaux sinistres incapacité S_x^N est donc la base des nouveaux sinistres de l'année N D_x^N où nous supprimons les contrats déjà sinistrés au 01/01/N de la base E_x^N soit

$$S_x^N = D_x^N - E_x^N$$

Notre base en N des valides V_x^N est donc la base des contrats en stock au 01/01/N A_x^N , des affaires nouvelles de l'année B_x^N , de l'ajout des contrats existants en sinistre mais pas en contrat F_x^N et de la suppression des contrats sinistrés au 01/01/N C_x^N soit

$$V_x^N = A_x^N + B_x^N + F_x^N - C_x^N$$

L'âge x est calculé de la façon suivante :

- Lorsque le contrat est sinistré, l'âge (arrondi à l'entier le plus proche) est l'âge à la date de survenance
- Lorsqu'il n'est pas sinistré, l'âge (arrondi à l'entier le plus proche) est l'âge au 01/01/N

Nous regroupons alors ces âges par tranches de 5 ans pour plus de fiabilité :

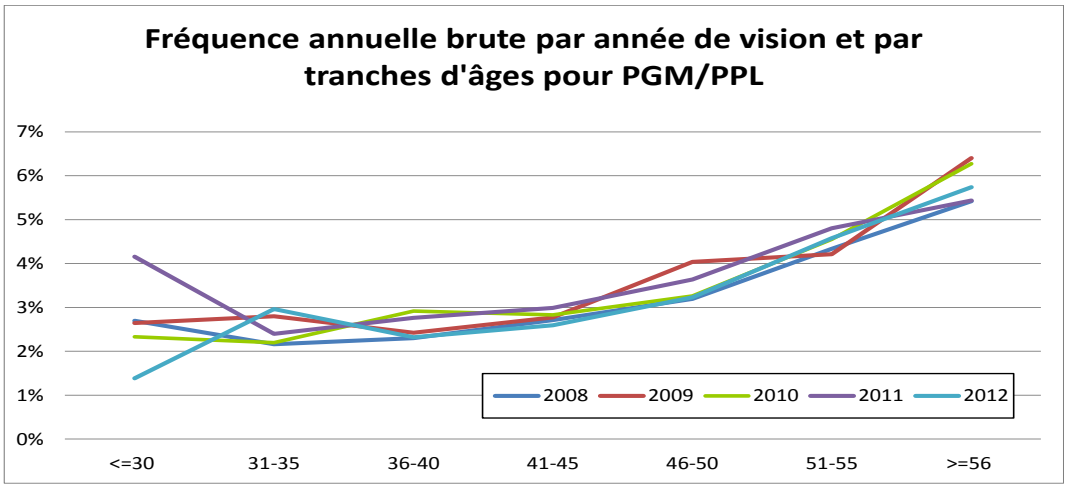
- $x \leq 30$
- $31 \leq x \leq 35$
- $36 \leq x \leq 40$
- $41 \leq x \leq 45$
- $46 \leq x \leq 50$
- $51 \leq x \leq 55$
- $56 \leq x$

Ce calcul d'effectifs est réalisé au 1er janvier de chaque année entre 2008 et 2012.

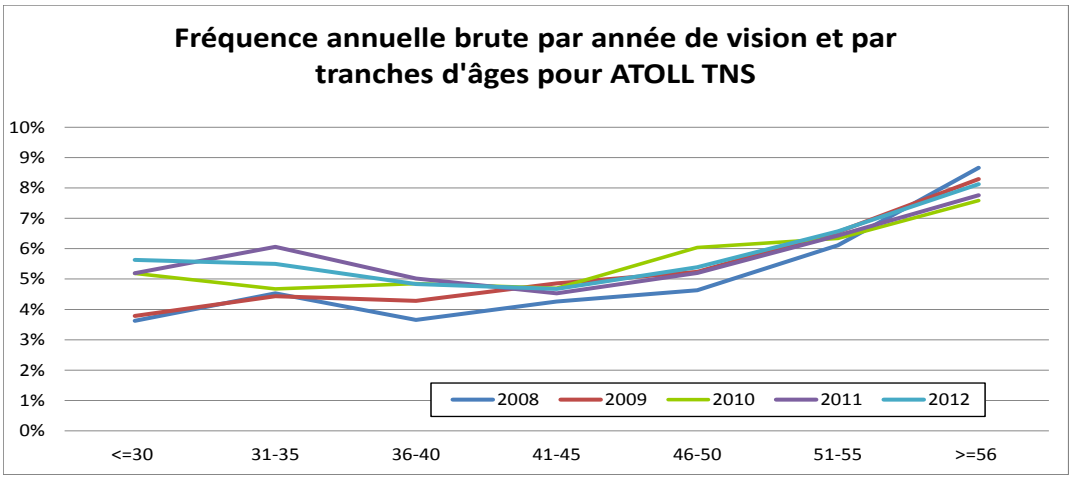
Le fichier final contient :

- le produit
- la classe d'âge
- l'année d'observation (ou année de vision)
- le nombre de valides
- le nombre de nouveaux sinistres en incapacité

Nous regardons par produit si l'année d'observation à un impact sur la fréquence observée.



Sur PGM/PPL, mise à part la tranche d'âges des moins de 30 ans où l'incidence varie de 1,4% à 4,2%, sur les autres classes, la fréquence observée sur les différentes années sont plutôt très proche. La volatilité de la première classe d'âges s'explique par un faible nombre de contrats exposé au risque.



Sur ATOLL, nous observons le même phénomène que PGM/PPL sur les premières classes d'âges. Cependant, aucune année ne se démarque de façon importante donc nous utiliserons toutes les années d'observation.

Nous sommons alors par produit et par classe d'âges le nombre de valides V_x et le nombre de nouveaux sinistres S_x quel que soit l'année de vision.
Pour obtenir notre fréquence brute moyenne nous appliquons la formule vue précédemment sur les nouveaux effectifs obtenus

$$Freq\ entrée\ incap_x = \frac{S_x}{V_x}$$

2. Le lissage

Sur ces taux bruts obtenus, nous appliquons un lissage afin d'obtenir des taux plus réguliers.

a. Méthode utilisée

La méthode choisie est une régression polynomiale par la méthode des moindres carrés afin d'obtenir une courbe qui s'ajuste au plus près de nos données.

L'équation de la droite polynomiale de degré p est la suivante :

$$\hat{y}_i = a_0 + a_1.x_i + a_2.x_i^2 + ... + a_p.x_i^p$$

Avec

- $X = (x_1, x_2, ..., x_n)$ un vecteur de n points représentant ici l'année de projection
- $Y = (y_1, y_2, ..., y_n)$ un vecteur de n points représentant les fréquences brutes que nous souhaitons lisser
- i = 1, 2, …, n avec $\hat{y}_i$ la valeur de y_i lissée

- a_i sont les coefficients à déterminer par résolution des équations :

$$\hat{Y} - Y = \begin{bmatrix} \hat{y}_1 - y_1 \\ \hat{y}_2 - y_2 \\ \cdots \\ \hat{y}_n - y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & .. & x_1^p & y_1 \\ 1 & x_2 & x_2^2 & \cdots & x_2^p & y_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 & \cdots & x_n^p & y_n \end{bmatrix} \times \begin{bmatrix} a_0 \\ a_1 \\ \cdots \\ a_p \\ -1 \end{bmatrix} = X \times A$$

Nous résolvons cette équation par la méthode des moindres carrés, c'est à dire que nous cherchons les coefficients $a_0, a_1, a_2 \ldots, a_p$ du polynôme qui minimisent :

$$P = \sum_{i=1}^{n} \left(\hat{y}_i - y_i\right)^2 = \left(\hat{Y} - Y\right)' \left(\hat{Y} - Y\right) = A'.X'.AX$$

Pour minimiser P, il faut annuler ses ($p+1$) dérivés partielles afin d'obtenir ($p+1$) équations à ($p+1$) inconnues $\frac{\partial(P)}{\partial a_k} = 0$ pour tout k = 0, 1, …, p.

Plus le degré du polynôme est élevé, plus la régression est précise, mais moins le lissage est réussi.

Pour valider le lissage et choisir le degré du polynôme, nous utilisons 2 méthodes :

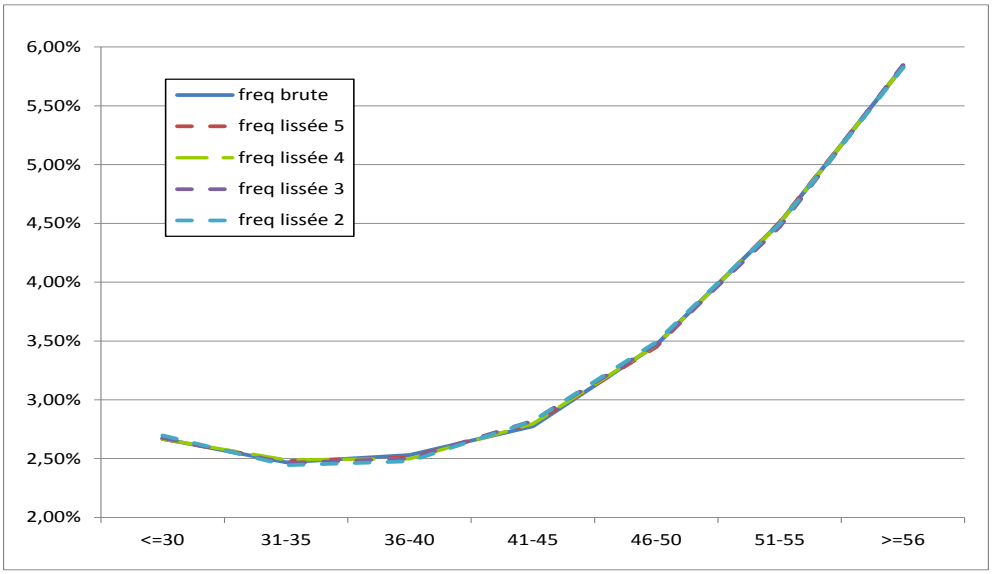
- la méthode graphique : nous observons les différentes courbes lissées et la courbe brute
- la méthode statistique : nous regardons la valeur du coefficient de détermination R^2. Cette statistique compare les valeurs estimées aux valeurs réelles et est comprise entre 0 et 1. Plus ce coefficient est proche de 1 et plus la corrélation est parfaite entre la courbe brute et la courbe lissée. A l'inverse, plus il est proche de 0 et moins la modélisation est satisfaisante. La formule de R^2 est la suivante : $R^2 = \frac{\sum_{i=1}^{n} \left(\hat{y}_i - \bar{y}\right)^2}{\sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2}$

b. Courbe lissée

Nous décidons de faire un lissage polynomial sur les âges médians pour chaque classe d'âge.
Nous lissons suivant plusieurs degrés du polynôme puis à l'aide des courbes graphiques et du R^2 nous choisissons pour chaque courbe le lissage le plus satisfaisant.

c. Choix du lissage

Pour PGM/PPL, voici le graphique des lissages testés suivant les différents degrés du polynôme :

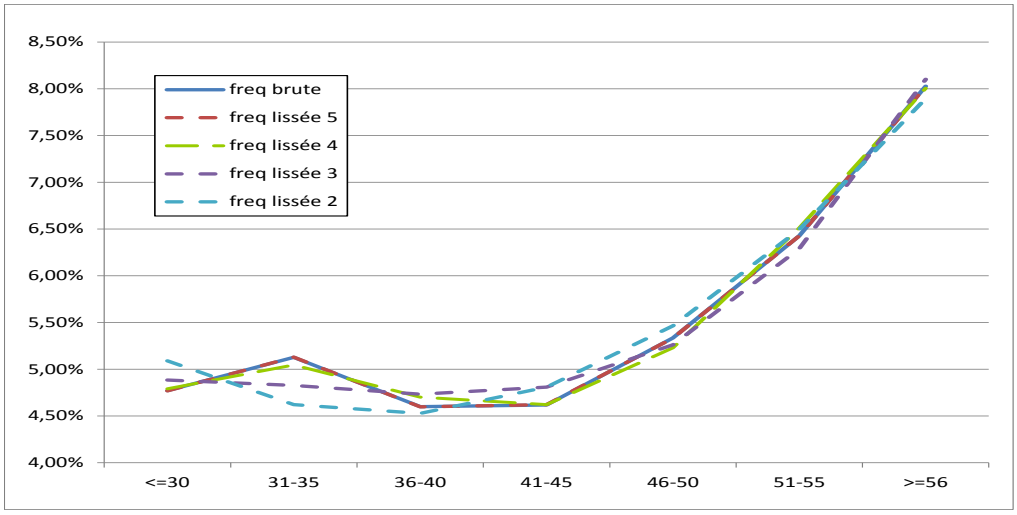


Visuellement, nous ne pouvons rien en conclure car toutes les courbes sont quasiment identiques.

	Polynome de degré 5	Polynome de degré 4	Polynome de degré 3	Polynome de degré 2
R^2	1,000	1,000	1,000	0,999

Statistiquement, nous pouvons conserver la régression de degré 2 qui est largement satisfaisante.

Voici également le graphique des différents lissages pour ATOLL TNS :



Graphiquement, nous observons que la régression polynomiale de degré 5 suit exactement la courbe brute donc ce lissage est sans intérêt. Le lissage de degré 2 est très éloigné de la courbe brute sur la tranche d'âge 31-35 ans.

	Polynome de degré 5	Polynome de degré 4	Polynome de degré 3	Polynome de degré 2
R^2	1,000	0,996	0,980	0,954

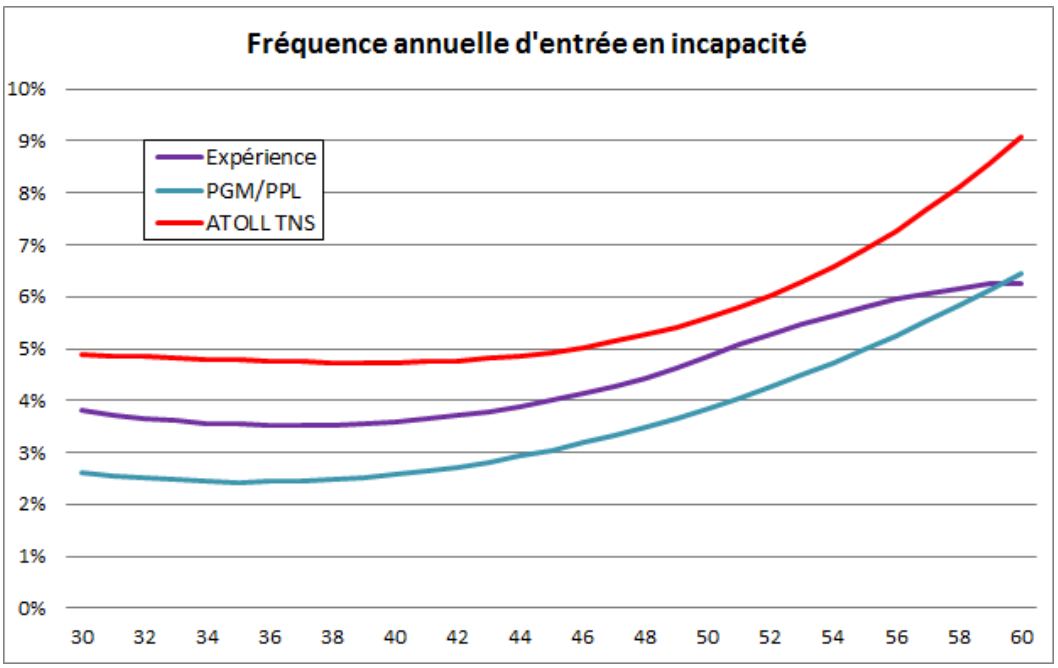
Le coefficient R^2 est très satisfaisant à partir de la régression de degré 3.

d. Reconstitution d'une courbe lissée par âge

A l'aide des coefficients de l'équation de la courbe lissée retenue, nous reconstruisons une fréquence d'entrée en incapacité par âge pour alimenter le modèle à la maille attendue. Au vue de la courbe obtenue pour les âges inférieurs à 30 ans, nous décidons d'appliquer la fréquence de 30 ans sur tous les âges inférieurs à 30 ans.

C. LES RESULTATS

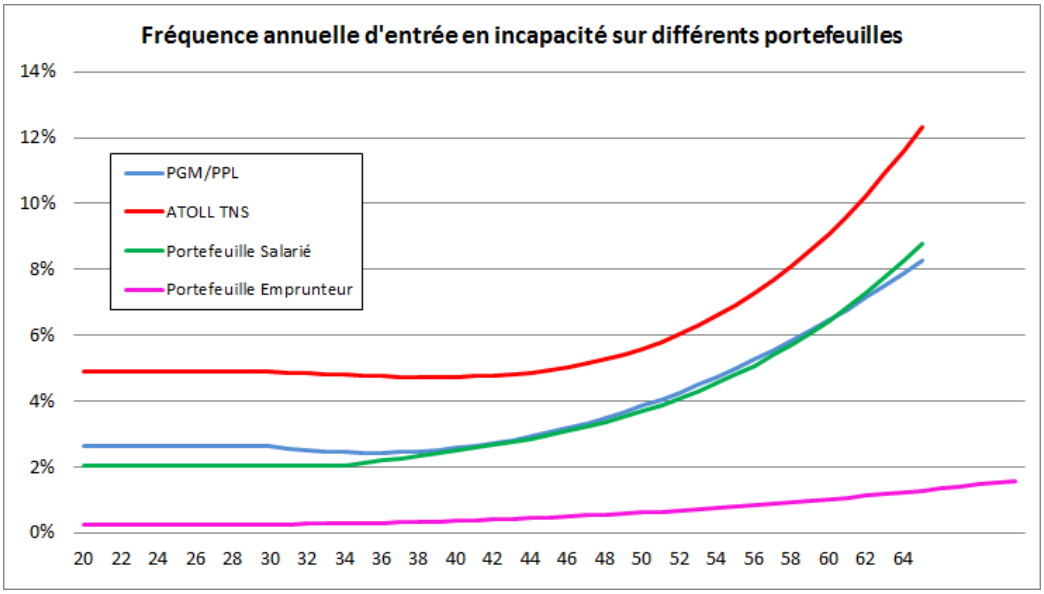
Voici les fréquences annuelles lissées obtenues par âges et par produits pour l'entrée en incapacité que nous comparons avec l'incidence déterminée par expérience en 2008 sur le portefeuille TNS:



Nous observons une tendance croissante en fonction de l'âge.

L'incidence en incapacité est clairement plus importante sur ATOLL TNS que sur PGM/PPL.
La fréquence obtenue lors de la construction des tables d'expérience 2008 se situe bien entre les courbes des 2 produits. Les allures générales des courbes sont les mêmes, modulo les âges supérieurs à 55 ans pour la fréquence d'expérience 2008.

Voici le graphique des lois d'incidence sur nos produits TNS que nous comparons aux fréquences d'entrée de nos produits salariés et emprunteurs de Generali (produits individuels) :

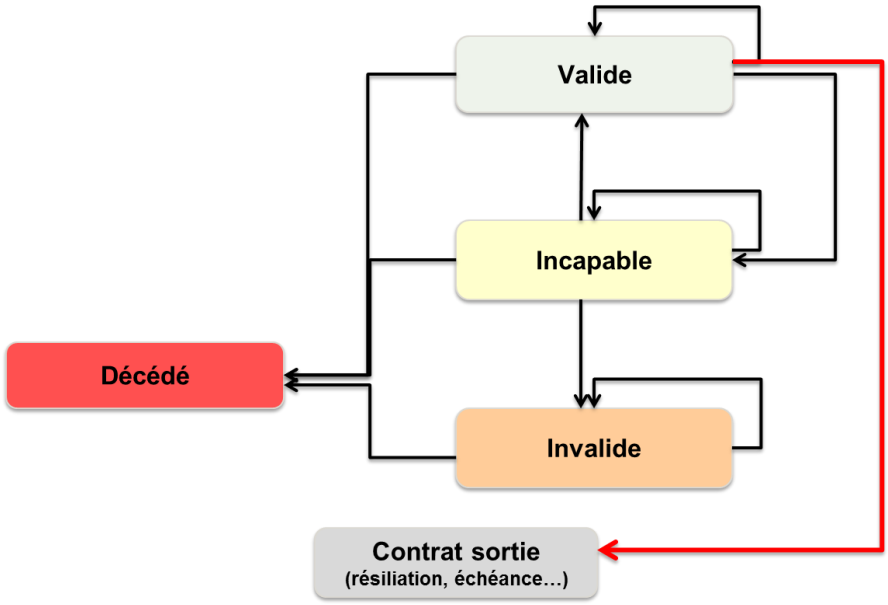


PGM/PPL se comporte plutôt comme un portefeuille de salariés alors qu'ATOLL a une incidence beaucoup plus importante.
La franchise a un impact important sur le taux d'incidence. En effet, plus la franchise est longue, plus la fréquence d'entrée en incapacité est faible, étant donné, qu'il faut dépasser la franchise pour être indemnisé. Ce qui explique la très faible incidence du portefeuille emprunteur ou quasiment toutes les franchises sont à 90 jours alors que les autres portefeuilles ont en majorité des franchises courtes (moins de 30 jours).
Il y a également l'antisélection à la souscription dans le choix de la franchise qui a un effet sur la fréquence d'entrée en incapacité. En effet, plus la franchise est courte, plus l'assuré va déclarer toute sorte de « petits » sinistres.

V. CONSTRUCTION DES LOIS DE CHUTES

Le but de cette partie est de décrire le processus appliqué pour la construction des lois de chutes des produits TNS, hypothèse nécessaire au modèle interne.

En effet, la projection du modèle commence à partir d'un stock de contrats à un instant donné qu'il faut faire « vivre » en simulant les sorties des contrats soit par décès (par une table de mortalité), soit par résiliation, soit par fin de garantie, par exemples, pour permettre la projection des primes futures. Pour Solvabilité 2, la projection de nos primes se fait jusqu'à la prochaine date anniversaire du contrat donc sur une durée maximale d'un an (principe de la frontière des contrats). En effet, nos contrats sont renouvelés par tacite reconduction et nous pouvons donc majorer chaque année les cotisations à la date anniversaire et l'assuré peut donc résilier son contrat tous les ans. Notre engagement s'arrête donc à l'anniversaire du contrat pour une vision Solvabilité 2 contrairement à la vision MCEV où les primes futures sont projetées jusqu'à extinction du portefeuille.



L'étude réalisée est basée sur un triangle avec projection par Chain-Ladder. La garantie (décès, incapacité, invalidité) n'est pas une variable retenue pour la segmentation de la loi; seul le produit et l'ancienneté segmentent la loi. En effet, par exemple lorsqu'un assuré souhaite résilier, il résilie en général, l'ensemble de ses garanties de son contrat. La garantie n'a pas d'impact sur la sortie du contrat.

Deux méthodes sont envisagées pour établir ces lois de chute :

- Triangle de nombre de contrats en cours
- Triangle de primes émises

Plusieurs difficultés ont été rencontrées concernant le triangle de primes :

- Les produits étudiés ayant migré fin 2008, l'historique actuellement disponible pour les primes n'est que de 5 ans (dans les fichiers récupérés sur l'ancien système, les primes ne sont pas présentes)
- Des résultats incohérents ont été obtenus notamment à cause des revalorisations et des majorations successives appliquées sur le montant des primes

La méthode par nombre a donc été privilégiée à la méthode par prime. Nous ne décrivons dans la suite, que le processus pour les triangles de nombre de contrats.

A. LES DONNEES

1. La présentation des bases de données

Plusieurs extractions vont servir à la construction de la base de données finale:

- Base 1 : une extraction à fin juin 2008 des contrats PGM/PPL et ATOLL du système de gestion d'origine. Ce fichier est donc la vision des contrats à un instant t.
- Base 2 : 11 fichiers spécialement extraits en 2012 du système d'information actuel pour la construction de ces lois de chutes : une vision des contrats à chaque fin d'année de 2000 à 2011
- Base 3 : la table SAS des contrats à fin décembre 2012 pour récupérer la vision à fin 2012 ; ici aussi, c'est une vision des contrats à un instant donné

En utilisant l'ensemble de ces extractions, nous récupérons ainsi l'historique complet du portefeuille de l'origine à fin 2012.

Toutes ces bases contiennent les informations essentielles à notre étude telles que :

- le numéro de contrat
- le produit
- la date d'effet du contrat (prise d'effet des garanties)
- la date d'émission du contrat (date à laquelle le contrat est émis dans le système)
- la date de sortie, renseignée si le contrat est effectivement sorti du portefeuille
- l'état du contrat à la date d'extraction (contrat en cours ou contrat sorti)

Ainsi, le taux de chute calculé représente la probabilité de sortie quel que soit la cause : décès, résiliation assuré, résiliation compagnie, échéance ou fin de garantie. Pour l'utilisation de ce taux dans le modèle, les décès en sont retirés car ils sont déjà modélisés par la loi de mortalité.

A noter, qu'il existe dans ces extractions une variable indiquant le motif de sortie (décès, sortie pour sinistre, résiliation pour impayés, résiliation volontaire, fin de contrat, sans effet….). Néanmoins, après analyse de cette donnée et rapprochement avec l'état du contrat, il semble que cette donnée ne soit pas fiable dans nos bases. Ce point a été remonté dans le cadre de la qualité

des données. Une fois ce problème traité, nous pourrions mettre à jour notre étude en distinguant les différentes causes de sortis.

2. Le traitement des données

Le traitement des données s'effectue sous SAS sur l'ensemble des fichiers sources.

a. Nettoyage des bases

Dans l'ensemble des bases, nous supprimons les contrats sans effet ou remplacés.

Sur la base 3, les contrats en cours sans date de passage (date liée au traitement informatique) sont des contrats en anomalies : en effet, ils sont sortis depuis plusieurs années et sont en erreur dans la table actuelle (après avoir remontée cette anomalie dans le cadre de la qualité de données, un plan de remédiation a été mis en place et le problème a été corrigé courant 2013). Nous supprimons donc ces contrats.

b. Traitement des bases

Pour la base 1, nous rapprochons l'ancien numéro de contrat avec le nouveau numéro de contrat correspondant quand il existe. Etant donné que, pour un contrat, nous avons une vision unique à une date t donnée, à l'aide de la date d'effet et de la date de sortie du contrat, nous reconstituons toutes les lignes par année de vision du contrat.

Par exemple, pour un contrat ayant une date d'effet au 01/01/2005 et une date de sortie au 15/05/2008 nous construisons 4 lignes :

- Année de vision = 2005 ; Etat du contrat = En cours
- Année de vision = 2006 ; Etat du contrat = En cours
- Année de vision = 2007 ; Etat du contrat = En cours
- Année de vision = 2008 ; Etat du contrat = Sorti

Nous agrégeons alors les tables par numéro de contrat, produit, état du contrat, année de vision. Pour un contrat et une année de vision, il n'y a donc qu'une seule ligne.

c. Consolidation des bases

Après homogénéisation des champs, nous consolidons les 3 bases sources avec les données suivantes:

- le numéro de contrat
- le produit
- l'année d'effet

- l'année de vision
- l'état du contrat à l'année d'observation
- la source (base1, base2 ou base3)

Nous créons une clé unique pour chaque ligne composée du numéro de contrat et de l'année de vision. Si un doublon est repéré, nous conservons la ligne issue du système de gestion actuel.

d. Retraitements des données

Dans le cadre de la QDD, voici quelques contrôles de cohérence que nous avons mis en place.

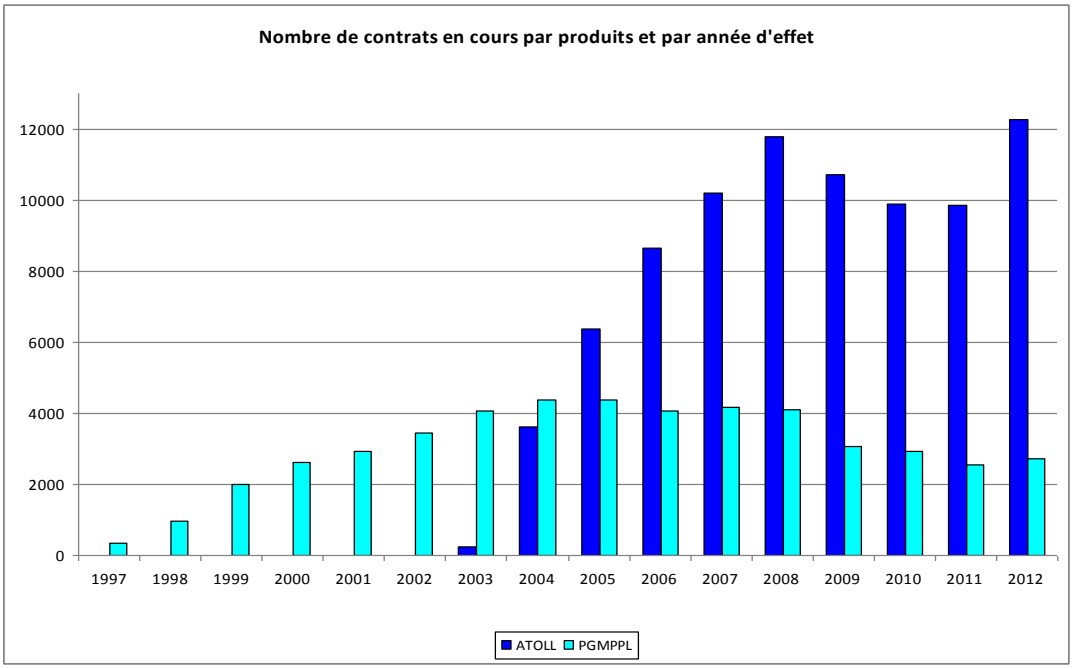
L'applicatif actuel gère l'effet retard : par exemple, un contrat peut être émis au 01/01/2013 avec une date d'effet au 01/12/2012. Ce contrat n'est donc connu qu'en 2013 mais il est garanti dès décembre 2012. Nous devons donc identifier ce contrat (en comparant la date d'effet et la date de mise en production) et rajouter une ligne avec une année d'observation à 2012 dans ce cas.

Certains contrats changent plusieurs fois de situation en passant de « sorti » à « en cours » (très souvent pour des annulations de résiliations pour impayés). Nous forçons donc l'état du contrat à « en cours » pour l'année de vision concernée. Nous créons ainsi une continuité dans l'état du contrat jusqu'à sa sortie définitive.

Une autre anomalie est apparue : il peut y avoir des « trous » d'observation c'est-à-dire qu'un contrat en cours va être observé sur 2008, 2009 et 2011. Pour qu'il y ait une bonne continuité de l'observation nous créons alors une ligne pour l'année de vision 2010. Ce problème peut être dû par exemple à un changement de situation temporaire : un contrat en cours peut passer à la situation sortie sur un mois pour non-paiement puis lors du règlement il redevient en cours.

3. Statistiques sur le portefeuille obtenu

Nous observons au total 49 295 contrats pour PGM/PPL, contrats durant en moyenne 5,5 ans. Pour ATOLL, nous avons 83 610 contrats ayant une durée moyenne de 3,5 ans.



Nous constatons que le produit PGM/PPL a mis quasiment 10 ans pour dépasser les 10 000 contrats souscrits par an, alors qu'ATOLL a très bien était souscrit dès le départ pour dépasser le nombre de contrats PGM/PPL en 2005 : ATOLL a une très forte croissance.

B. LA METHODE

1. Les taux bruts

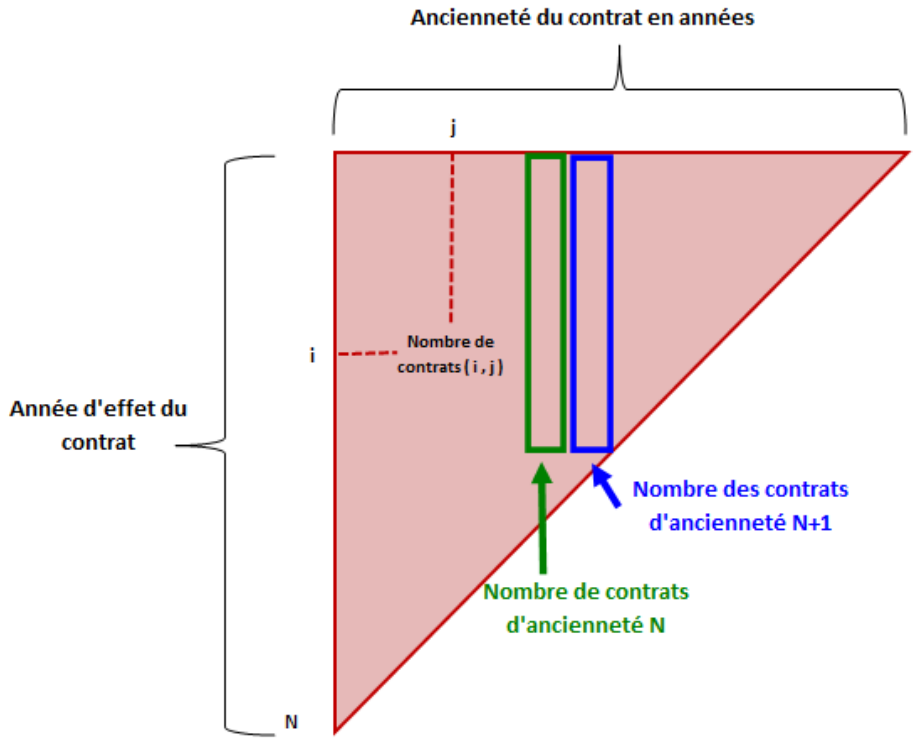
Nous utilisons des données propres à chaque portefeuille afin de tenir compte de la spécificité de chaque produit. Nous nous intéressons au comportement des contrats pour construire une loi de sortie suivant l'ancienneté du contrat en portefeuille.

Nous récupérons la table brut du nombre de contrats en cours par produit, année d'effet du contrat et année de vision.

Nous construisons un triangle de nombre de contrats en cours par année d'effet et durée du contrat. Ainsi, en sommant la dernière diagonale, nous avons le nombre de contrats en cours total en 2012 par exemple.
Un contrôle est mis en place pour vérifier que le nombre de contrats en cours pour une même année d'effet, diminue au fur et à mesure. Après l'analyse de plusieurs contrôles KO, nous avons identifié des données à retraiter (retraitements vus précédemment : contrat émis à effet rétroactif, trous d'observation pour un contrat ou contrat changeant plusieurs fois de situation).

Nous corrigeons donc ces anomalies et récupérons une base de nombre de contrats après correction.

Nous refaisons un triangle avec ces données corrigées :



Le nombre de contrats (i,j) représente le nombre de contrats en cours souscrit l'année i et ayant j années d'ancienneté.

Le taux de chute annuel du portefeuille entre une ancienneté N et N+1 est égal à :

$$\text{Taux de chute}_N = \frac{\text{Nombre de contrats sortis entre l'ancienneté N et N+1}}{\text{Nombre de contrats d'ancienneté N}}$$

$$= \frac{\text{Nombre de contrats d'ancienneté N - Nombre de contrats d'ancienneté N+1}}{\text{Nombre de contrats d'ancienneté N}}$$

$$= 1 - \frac{\text{Nombre de contrats d'ancienneté N+1}}{\text{Nombre de contrats d'ancienneté N}}$$

$$= 1 - f_j$$

Où f_j est le coefficient de développement déterminé par la méthode de Chain-Ladder :

$$f_j = \frac{\sum_{i=k-j}^{2012-j} \text{Nombre de contrats (i, j+1)}}{\sum_{i=k-j}^{2012-j} \text{Nombre de contrats (i, j)}}$$

avec k = 1989 pour PPL/PGM et k = 2004 pour ATOLL.

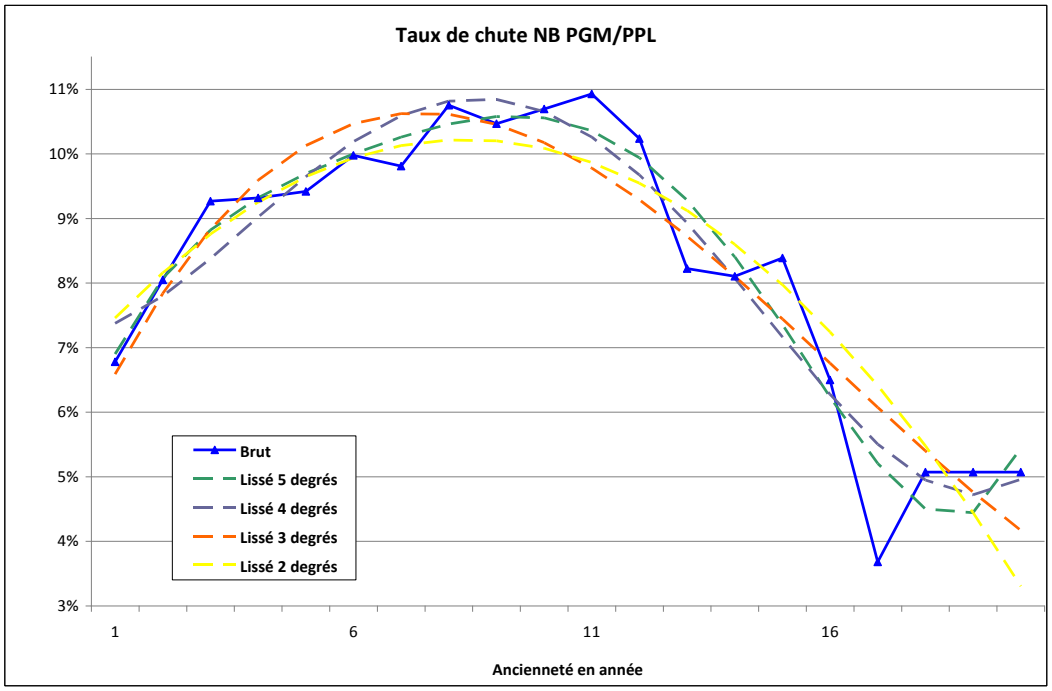
La méthode est appliquée tant que le nombre de contrats sur une même ancienneté dépasse les 100 observations. Etant donné la quasi stabilité des coefficients obtenus les dernières années, nous projetons le dernier coefficient jusqu'à la fin pour avoir une projection sur 20 ans.

Dans le modèle a états dans le cadre de Solvabilité 2, ce taux de chutes annuel est mensualisé pour faire sortir un peu de contrats tous les mois jusqu'à leurs dates anniversaires où à cet instant tous les contrats restants sortent.

2. Le lissage

Sur ces taux bruts obtenus, nous appliquons un lissage polynomial (méthode détaillé dans le chapitre précédent IV.B.2.a) afin d'obtenir des taux lissées.
Comme précédemment, pour choisir le degré du polynôme, nous utilisons la combinaison de la méthode graphique et de la méthode statistique.

a. Choix du lissage pour PGM/PPL



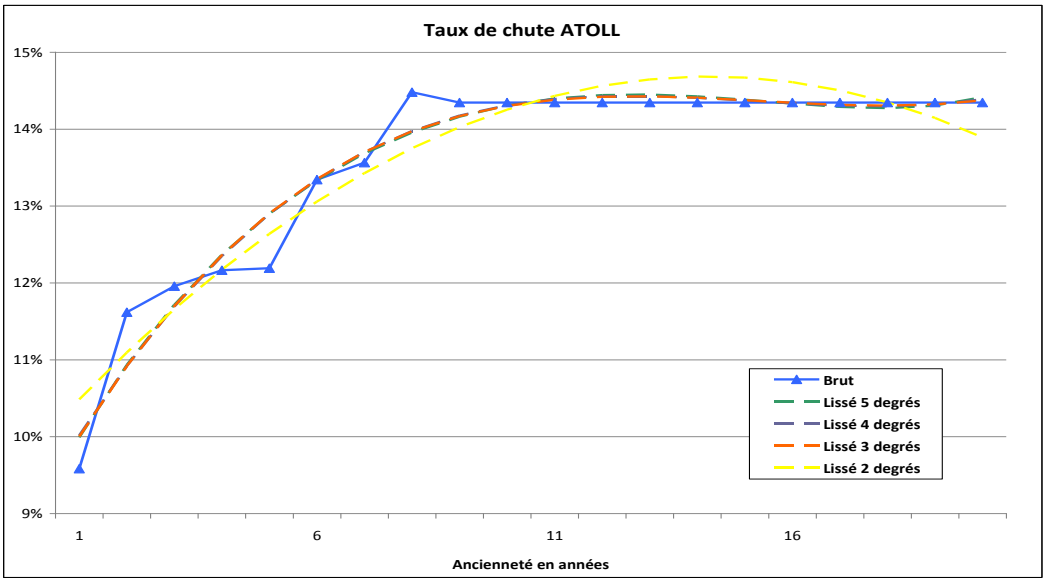
Visuellement, nous voyons clairement que la régression polynomiale de degré 2 ne convient pas car elle est trop éloignée du taux brut quasiment pour chaque âge. La régression polynomiale de degré 3 est écartées de la courbe brute à certains points, nous ne la retenons pas non plus.

	Polynome de degré 5	Polynome de degré 4	Polynome de degré 3	Polynome de degré 2
R^2	0,93	0,91	0,87	0,83

Statistiquement, nous pouvons conserver les régressions de degré 5.

Au vu de ces deux éléments, nous choisissons la régression de degré 5.

b. Choix du lissage pour ATOLL



Visuellement, nous voyons que la régression polynomiale de degré 2 ne convient pas car éloignée de la courbe brute à de nombreux points.

	Polynome de degré 5	Polynome de degré 4	Polynome de degré 3	Polynome de degré 2
R^2	0,953	0,953	0,953	0,918

Le coefficient R^2 est très satisfaisant à partir de la régression de degré 3.

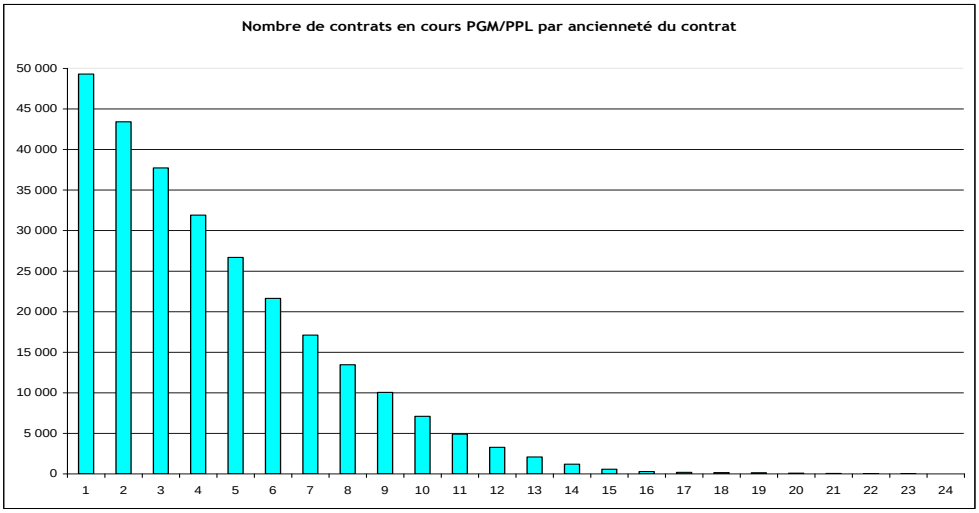
Au final, nous choisissons la courbe lissée à l'aide du polynôme de degré 3.

C. LES RESULTATS

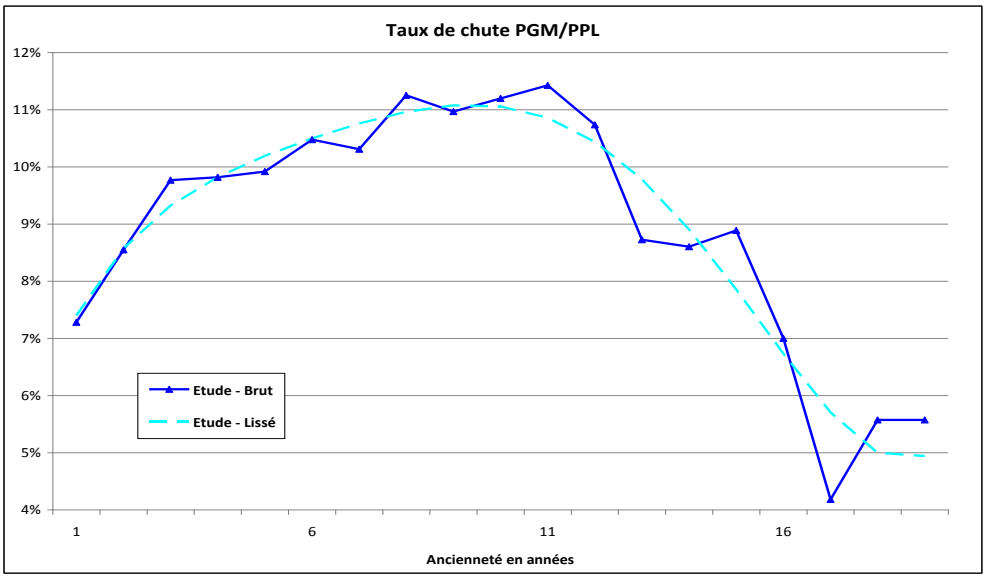
Nous comparons graphiquement la courbe brute et la courbe lissée par année de projection par produit.

1. Résultats PGM/PPL

Sur PGM/PPL, nous avons récupéré les données des deux systèmes de gestion, donc notre base de données est complète depuis la création du produit en 1989 soit un recul de 24 ans.



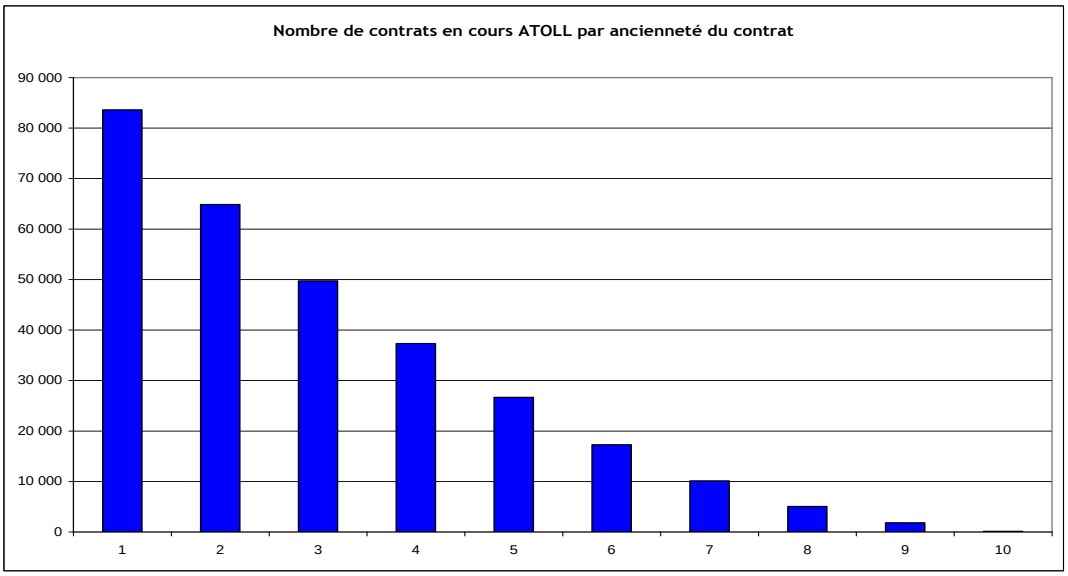
A noter, cependant que la commercialisation de PGM/PPL n'a pas été importante les 10 premières années et nous n'avons donc pas beaucoup de données sur les anciennetés importantes.

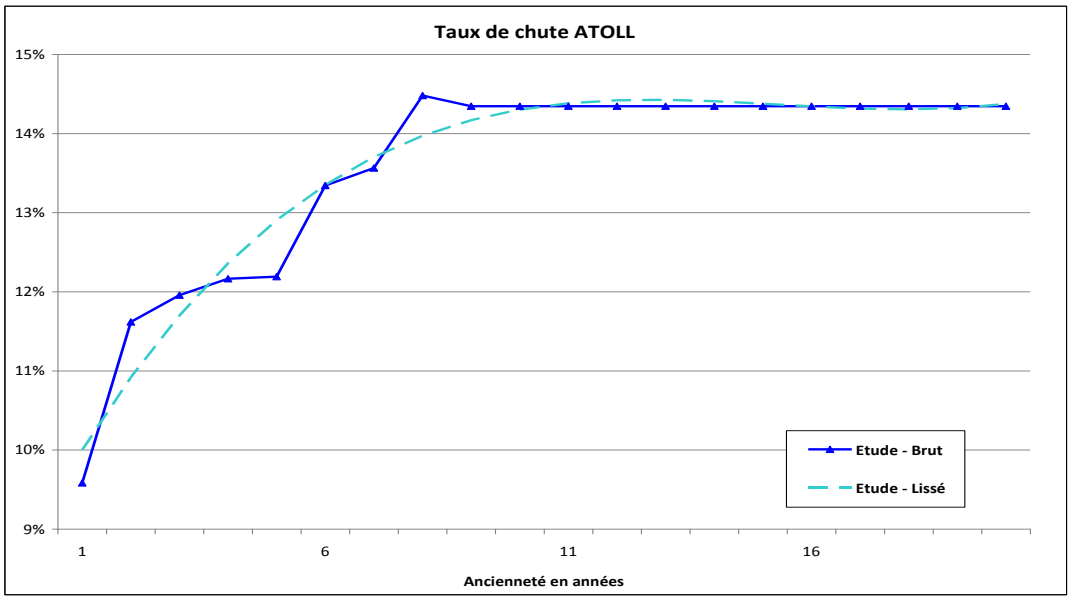


A partir de 18 ans de projection, le nombre de contrats n'étant plus significatif, le coefficient de Chain-Ladder de la 18ème année est reporté pour les années suivantes; d'où la stabilité observée. Nous constatons une augmentation des chutes pendant les 10 premières années de souscription puis ce taux diminue progressivement.

2. Résultats ATOLL TNS

Sur ATOLL nous avons récupéré l'intégralité des données donc notre base de données est complète depuis la création du produit en 2004 soit un recul de 9 ans.





A partir de 9 ans de projection, le nombre de contrats n'étant plus significatif, le coefficient de Chain-Ladder de la 9ème année est reporté pour les années suivantes; d'où la stabilité observée.

Le produit ATOLL chute plus que le produit PGM/PPL et le taux de sortie ne fait qu'augmenter.
A noter que cette différence de comportement entre ATOLL et PGM/PPL est certainement expliquée par les majorations de cotisations importantes et successives (tous les ans) appliquées à ATOLL notamment sur les garanties incapacité/invalidité. Ces majorations concernent aussi bien les affaires nouvelles que le stock de contrats en cours. En général, les augmentations de PGM/PPL sont beaucoup plus faibles que celles d'ATOLL.
La conséquence de ces importantes majorations est la résiliation de contrats par les assurés trouvant des tarifs plus attractifs chez nos concurrents. Ces assurés sont donc « plutôt des bons risques ». Les assurés ne résiliant pas leur contrat sont des « mauvais risques »; ils ne trouvent pas un meilleur tarif chez les autres assureurs, soit pour une question d'âge, soit de problèmes de santé, soit de prestations garanties très importantes…

CONCLUSION

L'étude réalisée s'inscrit dans le cadre de Solvabilité 2 et du modèle interne d'un portefeuille de prévoyance TNS. Pour alimenter le modèle à états, nous avons besoin de diverses hypothèses réelles liées à ce portefeuille spécifique et notamment pour la garantie incapacité/invalidité.

A l'aide de quelques statistiques comparant notre portefeuille TNS et notre portefeuille salariés, nous remarquons des différences entre ces deux populations concernant la souscription (pour les TNS, les montants des indemnités garanties sont beaucoup plus élevés, les franchises sont plus courtes) et la sinistralité (lorsque les TNS sont sinistrés, ils reprennent plus vite leur activité).

Concernant le retraitement des données, nous en avons contrôlé la cohérence, l'exhaustivité et l'exactitude comme exigé par Solvabilité 2. Selon le type d'anomalie constatée, nous avons choisi entre :

- Corriger automatiquement la valeur fausse ou manquante en la remplaçant par une valeur potentiellement correcte
- Procéder à une correction approximative à l'aide d'autres variables
- Agréger plusieurs sinistres en un seul (si rechute non identifiée par exemple)
- Supprimer le sinistre contenant la valeur aberrante si une valeur correcte n'a pas pu être identifiée

Nous pouvons considérer que la qualité des données est correcte car moins de 6% des données n'ont pu être utilisées.

Nous avons ainsi construit des tables d'expérience pour la loi de maintien en incapacité et le passage en invalidité. Nous avons utilisé la méthode de l'estimateur de Kaplan Meier ainsi que le lissage de Whittaker Henderson. Le nombre de sinistres invalidité étant trop faible pour construire une table de maintien en invalidité robuste, pour calculer l'impact sur les provisions, nous avons utilisé la table créée par le BCAC en cours de test chez les différents assureurs. En effet, cette nouvelle table reflète plus la sinistralité réelle que la table règlementaire datant de 1996.
Nous observons une baisse totale de 3% sur l'ensemble des provisions de notre portefeuille TNS en utilisant les tables de notre étude par rapport aux tables utilisées aujourd'hui dans les comptes (lois d'expérience 2008 pour l'incapacité et loi règlementaire pour l'invalidité).
Les comparaisons effectuées sur les différentes lois (lois règlementaires du BCAC, lois d'expérience 2008 et lois issues de notre étude) et sur les montants des provisions sont intéressantes car elles montrent que :

- la sinistralité propre à notre portefeuille diverge largement de celle établit sur une population de salariés du BCAC
- un ajustement était nécessaire par rapport aux tables construites en 2008 pour calculer des provisions en Best Estimate et pour une modélisation plus fine

Cela justifie donc notre choix de vouloir construire des tables d'expérience sur notre propre portefeuille d'assurés.

Nous avons également établi une loi d'incidence en incapacité par produit et par tranche d'âges, basée sur l'observation annuelle des nouveaux incapables rapportés à l'exposition. En comparant les résultats obtenus avec différents portefeuilles (salariés, emprunteurs de Generali), nous observons une incidence différente sur les deux produits TNS et un impact important des franchises et de l'antisélection.

Enfin, nous avons mis en place une loi de chutes à partir de la méthode Chain-Ladder appliquée aux triangles de nombres de contrats en cours. Sur cette loi également, le comportement des assurés des deux produits étudiés est différent : sur un produit le taux de chute est plus élevé certainement à cause des majorations de cotisations successives et importantes qui lui sont appliquées.

Tout au long de notre étude, nous avons été confrontés aux problématiques de qualité de données. Nous avons détecté les anomalies les plus importantes pour les corriger afin d'obtenir une base de travail la plus propre possible. Ces problèmes de données sont causés :

- soit par l'environnement informatique qui dégrade la qualité des données comme des migrations réalisées en prenant des raccourcis par manque de temps et de budget sans discuter des impacts éventuels avec les différents métiers
- soit par les systèmes de gestion ayant des fonctionnements compliqués (gestion de deux dates de survenance par exemple) ou ne bloquant aucune saisie incohérente, laissant une liberté totale aux gestionnaires
- soit par les systèmes décisionnels ne reflétant pas la réalité de l'applicatif de gestion
- soit par l'homme directement par des erreurs de saisies

Le travail sur la qualité des données effectué pour cette étude, nous a permis de mettre en place des plans de remédiation avec l'informatique et la gestion afin d'améliorer notre qualité des données et de sensibiliser ces départements à l'importance d'avoir des données correctes pour effectuer nos études.

L'ensemble de nos tables construites servent ainsi à modéliser le risque incapacité/invalidité du portefeuille TNS de Generali :

- en calculant les provisions en Best-Estimate
- en projetant les cash-flows
- en alimentant le modèle interne des probabilités de passage d'un état à un autre à chaque pas de projection
- en permettant de calibrer les chocs à appliquer sur la fréquence et la sévérité de la sinistralité pour simuler les différents scénarios possibles pour le calcul du SCR

ANNEXES

Annexe 1 : Graphiques des intervalles de confiance pour la loi brute de maintien en incapacité

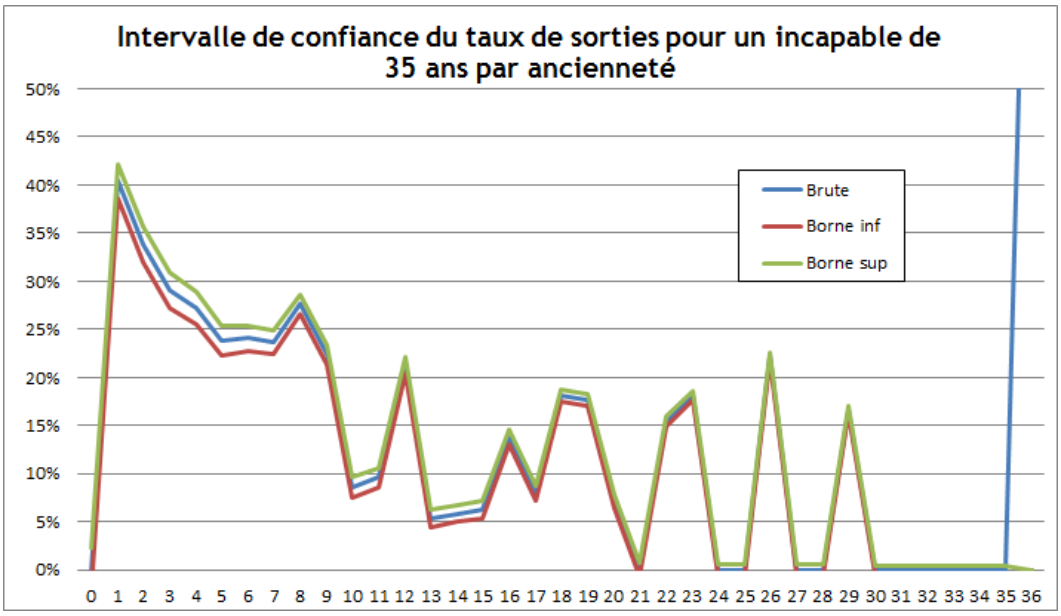
Annexe 2 : Méthode du lissage de Whittaker Henderson

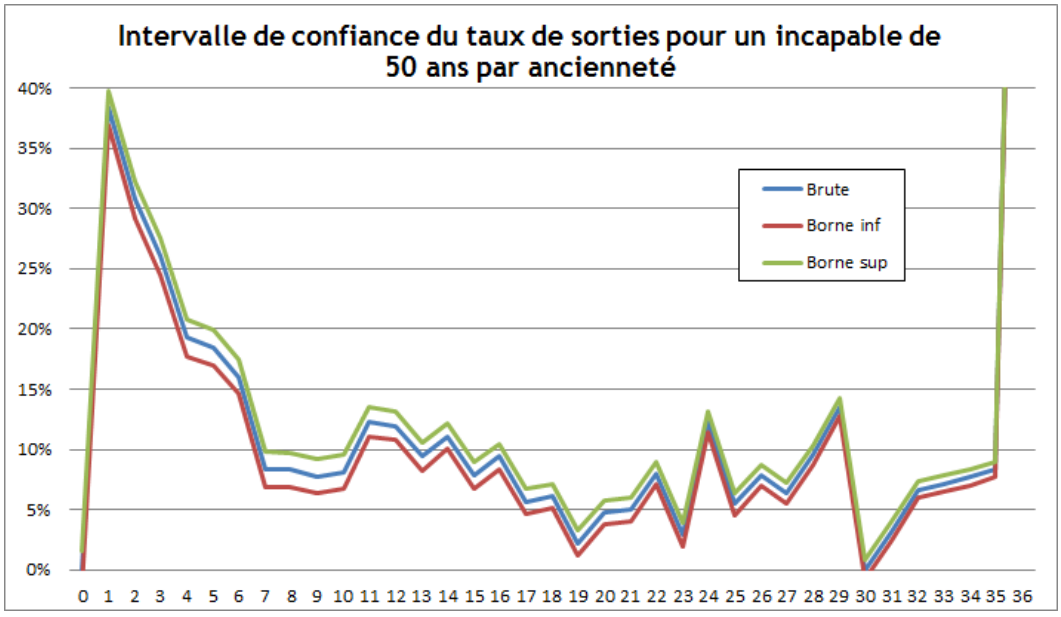
Annexe 3 : Comparaisons des tables lissées de maintien en incapacité avec le BCAC et les anciennes tables d'expérience

Annexe 4 a : Comparaisons des tables lissées de passages en invalidité avec le BCAC et les tables d'expérience 2008

Annexe 4 b : Comparaisons des nombres de passages cumulés par ancienneté entre notre étude, le BCAC et les tables d'expérience 2008

Annexe 1 : Graphiques des intervalles de confiance pour la loi brute de maintien en incapacité





Annexe 2 : Méthode du lissage de Whittaker Henderson

Notons u_1, u_2, u_3, …, u_n les valeurs observées et v_1, v_2, v_3, …, v_n les valeurs lissées.

i. Le critère de fidélité

Le critère de fidélité mesurant la distance euclidienne classique entre la valeur lissée v_i et la valeur observée u_i quantifie la qualité de l'ajustement. Il est définit par :

$$F = \sum_{i=1}^{n} w_i \left(v_i - u_i \right)^2$$

avec w_i le poids affecté au respect de la donnée i

n l'âge maximum d'incapacité

Ainsi, si les données lissées sont égales aux données brutes alors F est nul.

ii. Le critère de régularité

Le critère de régularité mesurant la régularité de la courbe lissé est définit par la formule suivante :

$$S = \sum_{i=1}^{n-Z} (\Delta^Z v_i)^2$$

avec Z-1 le degré du polynôme utilisé pour la régularité (Z est donc un entier positif)

$\Delta^Z v$ la différence d'ordre z entre les valeurs du vecteur v, ainsi nous avons :

$$\Delta v(x) = v(x+1) - v(x) \text{ et } \Delta^2 v(x) = \Delta(\Delta v(x)) = v(x+2) - 2v(x+1) + v(x) \ldots$$

Ainsi, si les données lissées suivent exactement un polynôme de degré Z-1, S sera nul car $\Delta^z v_i$ =0.

iii. La résolution de la minimisation

La méthode de Whittaker-Henderson consiste à trouver les valeurs lissées v_i qui minimisent M définit par une combinaison linéaire de la fidélité et de la régularité : $M = F + hS$ avec h un paramètre d'arbitrage entre fidélité et régularité.

Il faut donc résoudre le système des n équations suivantes :

$$\forall i, \frac{\partial M}{\partial v_i} = 0$$

Une approche matricielle permet de résoudre ce problème de minimisation.

Posons les notations suivantes :

- $U^t = [u_1, u_2, \ldots, u_n]$
- $V^t = [v_1, v_2, \ldots, v_n]$
- $W = \begin{pmatrix} w_1 & & 0 \\ & \ddots & \\ 0 & & w_n \end{pmatrix}$
- K_z la matrice contenant les coefficients binomiaux d'ordre z avec alternance de leur signe tels que:

$\Delta^z V = Kz.V = \begin{bmatrix} \Delta^z v_1 \\ \Delta^z v_2 \\ \ldots \\ \ldots \\ \Delta^z v_{n-z} \end{bmatrix}$ avec K_z de dimension (n-z)*n

Par exemple : $K_2 = \begin{pmatrix} 1 & -2 & 1 & 0 & & & 0 \\ 0 & \ddots & \ddots & \ddots & & & \\ & & & \ddots & \ddots & \ddots & 0 \\ 0 & & & 0 & 1 & -2 & 1 \end{pmatrix}$

Nous injectons ces formules dans l'expression de base de M :

$$M = (V-U)^t W (V-U) + h(K_z V)^t K_z V$$
$$M = (V-U)^t W (V-U) + hV^t \left(K_z{}^t K_z\right) V$$
$$M = V^t WV - V^t WU - U^t WV + U^t WU + hV^t \left(K_z{}^t K_z\right) V$$
$$M = V^t WV - 2V^t WU + U^t WU + hV^t \left(K_z{}^t K_z\right) V$$

Dérivons M : $\frac{\partial M}{\partial V} = 2WV - 2WU + 2hK_z^t K_z V$ car W est symétrique

La résolution de $\frac{\partial M}{\partial V} = 0$ conduit alors à l'expression : $WV + hK_z^t K_z V = WU$

Si nous posons $C = W + hK_Z^t K_Z$ alors nous obtenons $CV = WU$.

C est une matrice symétrique, montrons que C est une matrice définie positive :

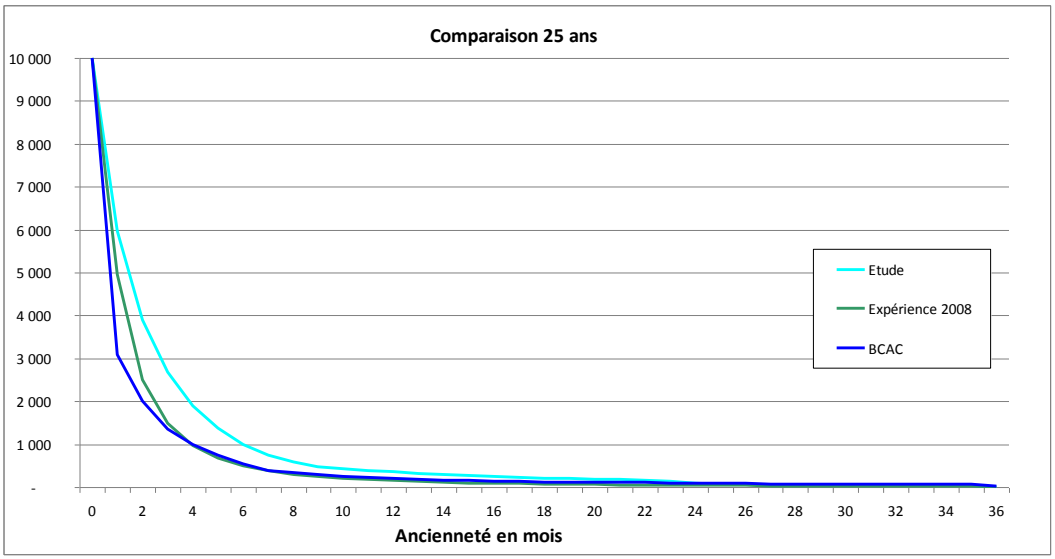
Soit Y un vecteur colonne de taille n, $Y^t CY = \sum_{i=1}^{n} w_i {y_i}^2 + h\sum_{i=1}^{n-z} \left(\Delta^z y_i\right)^2$

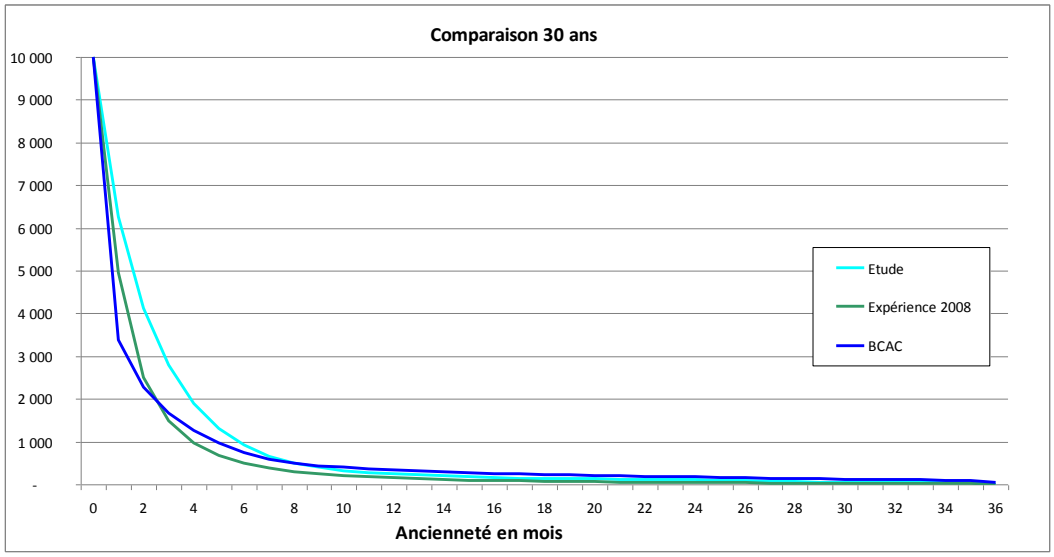
Nous en déduisons le résultat suivant : $Y^t CY \geq 0$ et $Y^t CY = 0 \Rightarrow Y = 0$

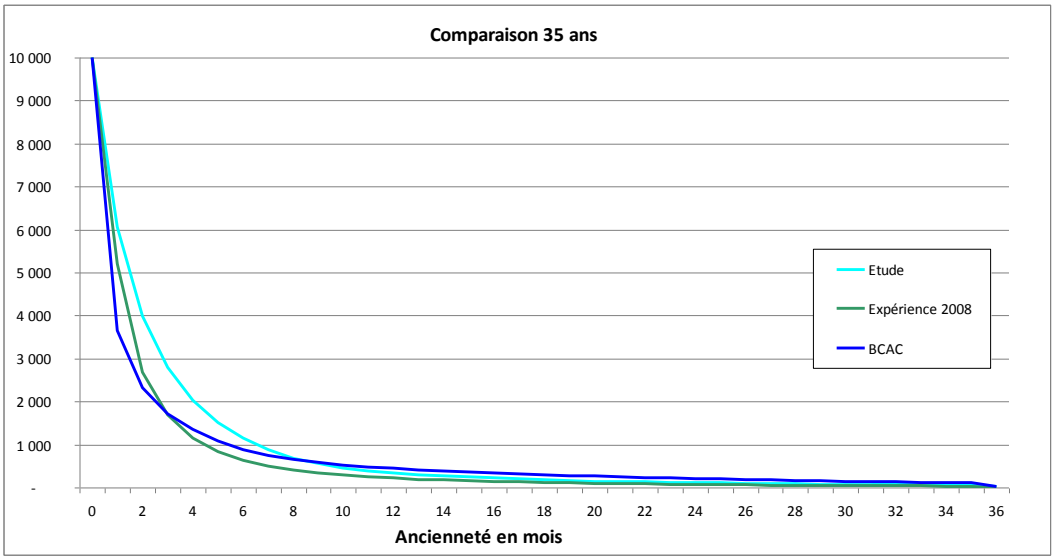
C est donc symétrique et définie positive, elle est donc inversible.

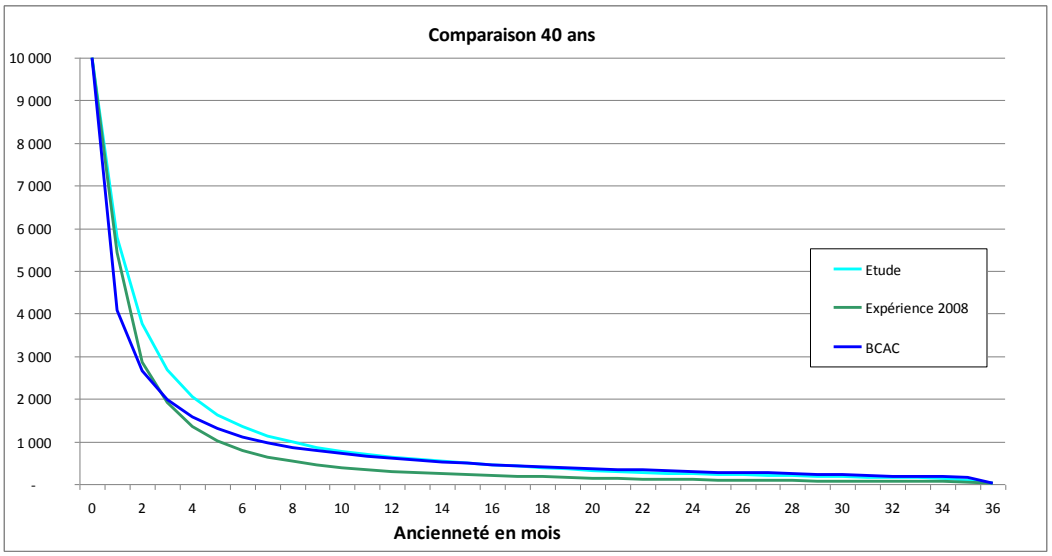
Nous obtenons ainsi la solution suivante : $V = C^{-1}WU$

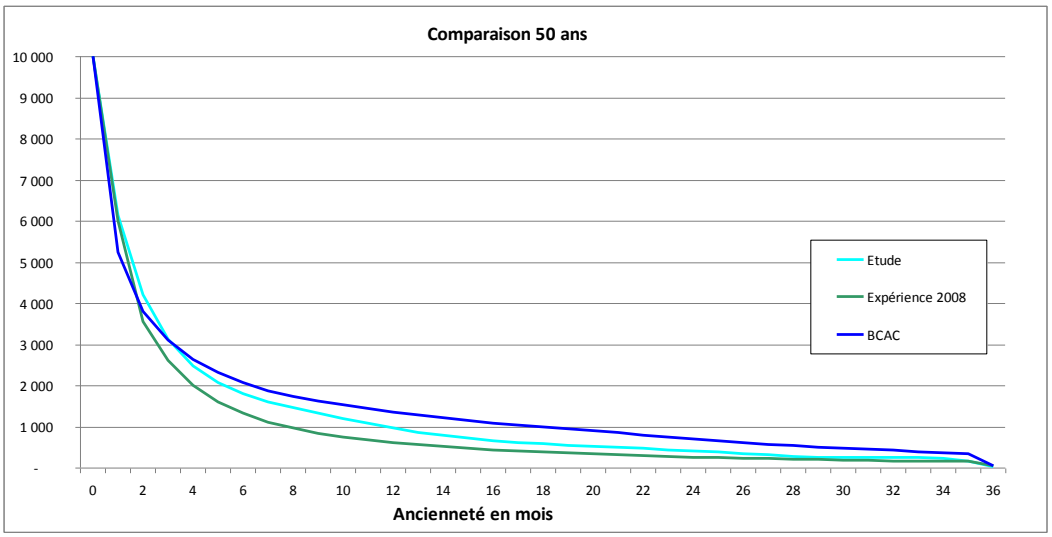
Annexe 3 : Comparaisons des tables lissées de maintien en incapacité avec le BCAC et les anciennes tables d'expérience

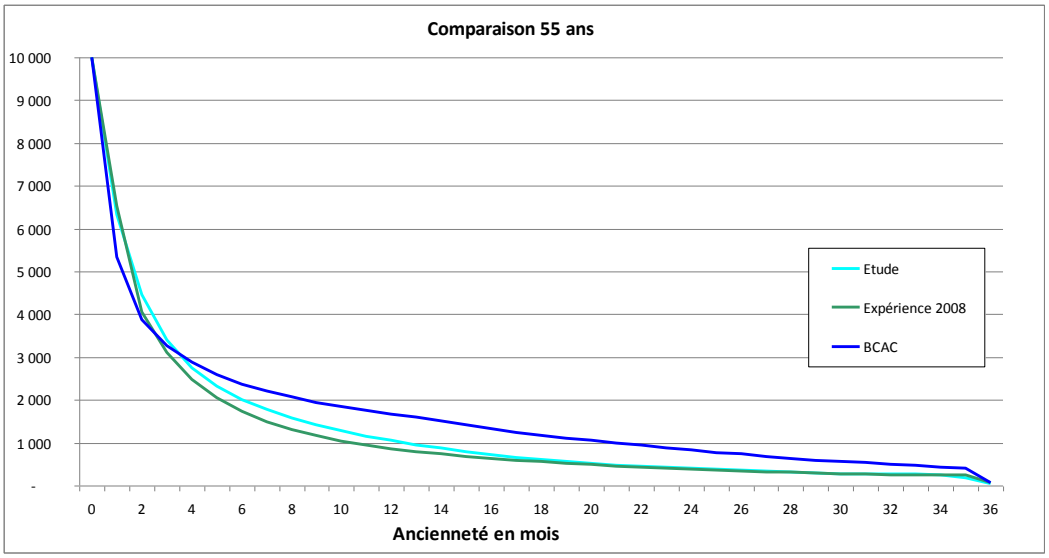




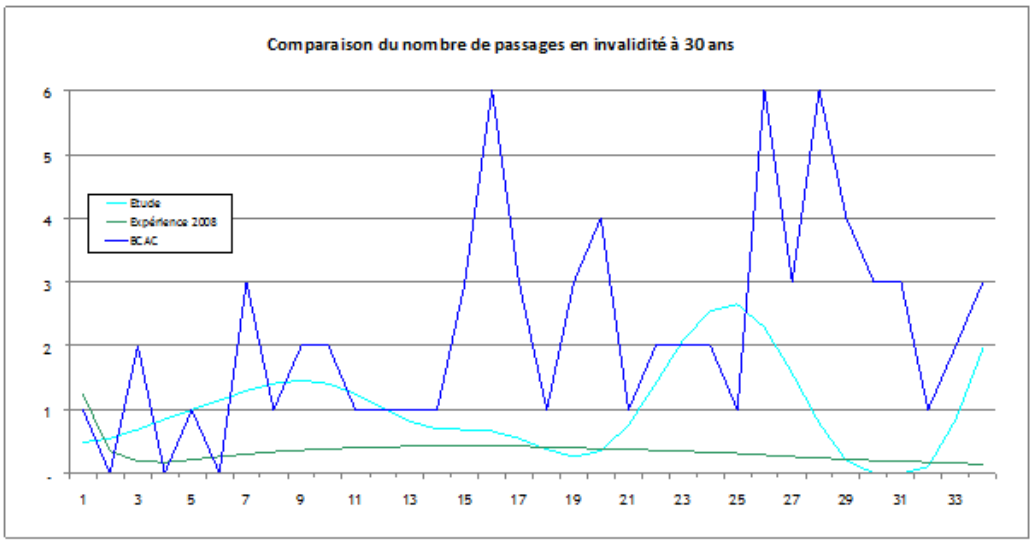


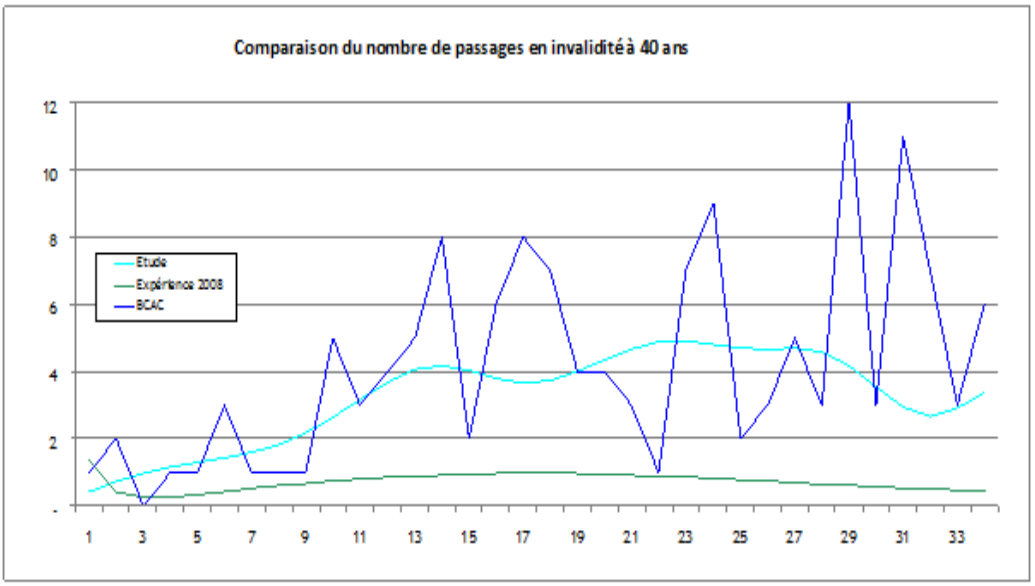

Comparaison 40 ans
10 000
9 000
8 000
7 000
6 000
5 000
4 000
3 000
2 000
1 000
-
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36
Ancienneté en mois
Etude
Expérience 2008
BCAC

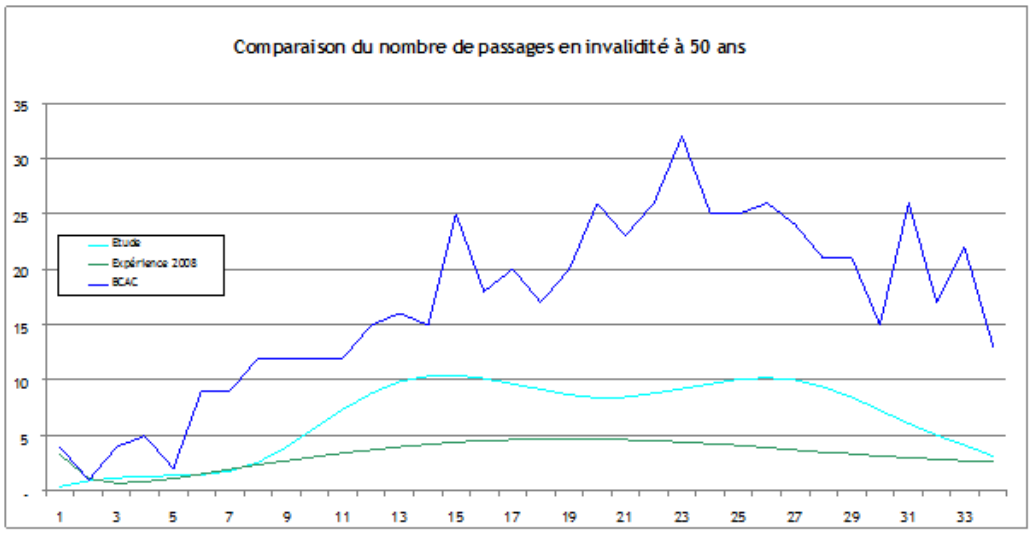

Comparaison 50 ans
10 000
9 000
8 000
7 000
6 000
5 000
4 000
3 000
2 000
1 000
-
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36
Ancienneté en mois
Etude
Expérience 2008
BCAC

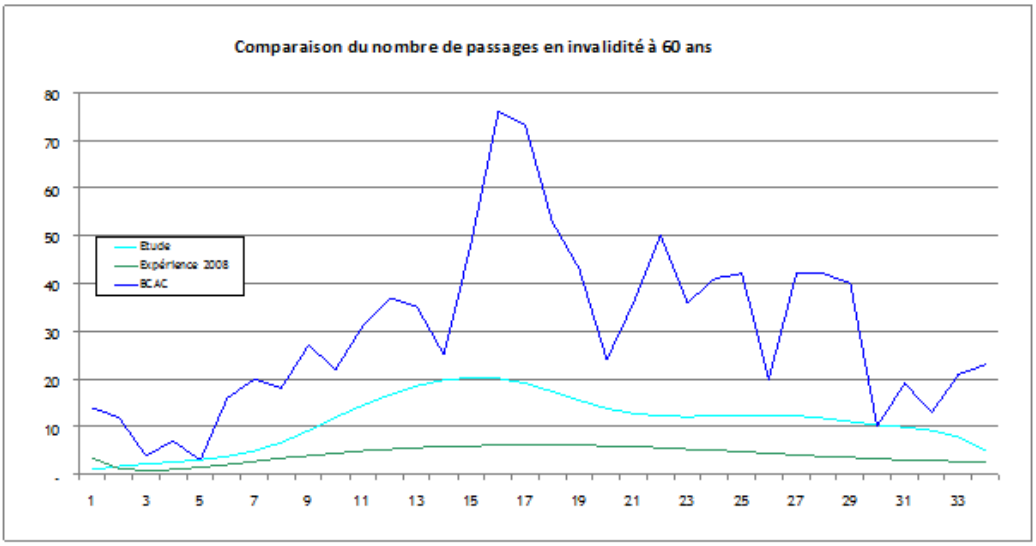

Comparaison 55 ans
10 000
9 000
8 000
7 000
6 000
5 000
4 000
3 000
2 000
1 000
-
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36
Ancienneté en mois
Etude
Expérience 2008
BCAC

Annexe 4 a : Comparaisons des tables lissées de passages en invalidité avec le BCAC et les tables d'expérience 2008

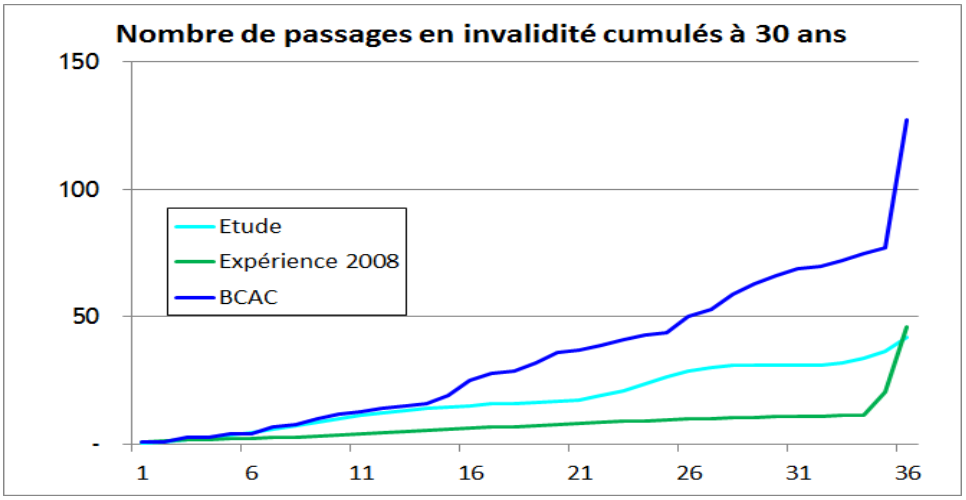


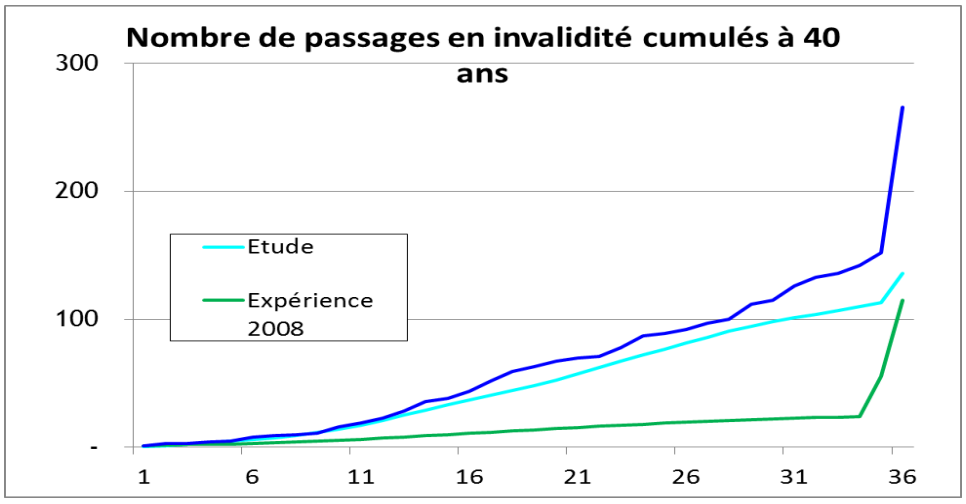


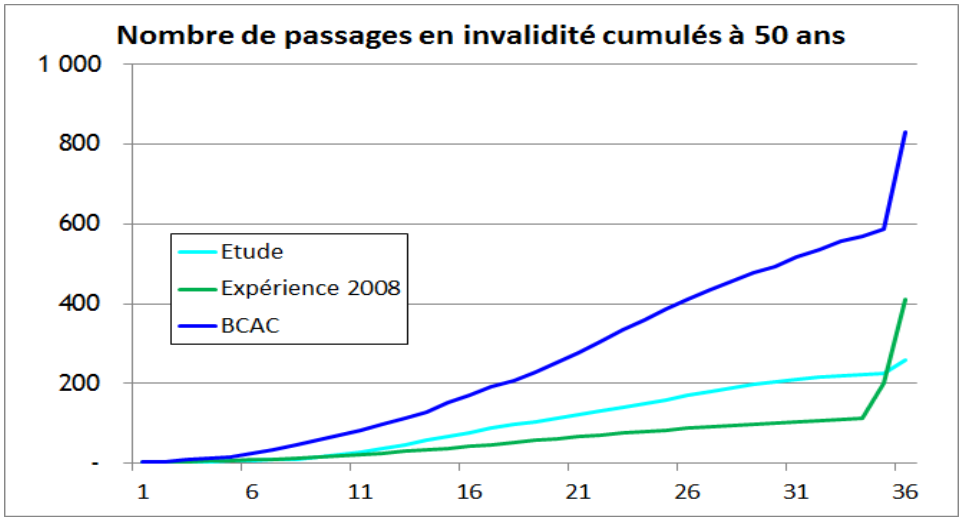

Comparaison du nombre de passages en invalidité à 50 ans
35
30
25
20
15
10
5
-
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
Etude
Expérience 2008
BCAC

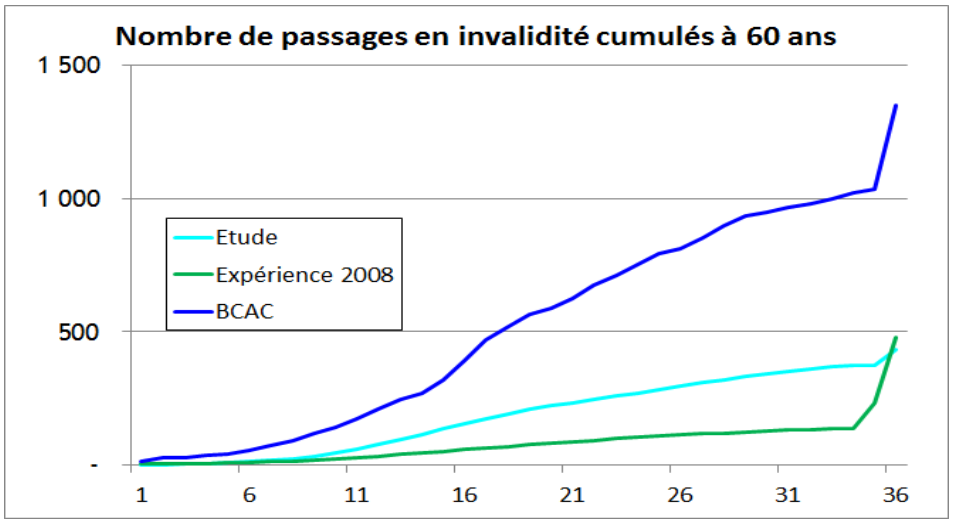

Comparaison du nombre de passages en invalidité à 60 ans
80
70
60
50
40
30
20
10
-
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
Etude
Expérience 2008
BCAC

Annexe 4 b : Comparaisons des nombres de passages cumulés par ancienneté entre notre étude, le BCAC et les tables d'expérience 2008






Nombre de passages en invalidité cumulés à 50 ans
1 000
800
600
400
200
-
1
6
11
16
21
26
31
36
Etude
Expérience 2008
BCAC


Nombre de passages en invalidité cumulés à 60 ans
1 500
1 000
500
-
1
6
11
16
21
26
31
36
Etude
Expérience 2008
BCAC

BIBLIOGRAPHIE

Sites internet

- FFSA : www.ffsa.fr
- Ressources actuarielles : www.ressources-actuarielles.net
- Actu d'actuaires : www.actudactuaires.typepad.com
- Code des assurances, code de la sécurité sociale… : www.legifrance.gouv.fr
- Sécurité sociale des indépendants et des salariés : www.rsi.fr et www.amelie.fr
- Institut National de la Statistique et des Etudes Economiques : www.insee.fr
- Autorité de Contrôle Prudentiel et de Résolution : www.acpr.banque-france.fr
- European Insurance and Occupational Pensions Authority : www.eiopa.europa.eu

Documentations internes

- Documentation sur la construction des lois d'expérience de 2008
- Conditions Générales des produits PGM/PPL et ATOLL
- Présentations sur la QDD

Articles

- « Lignes directrices de la construction des lois de maintien en incapacité et en invalidité », Igotz Aubin et Anne Rolland, IA 2010
- « Refonte des Lois de Maintien en Incapacité temporaire », Note technique BCAC, avril 2013
- « Les contrats prévoyance Madelin en 2013 », Enquête FFSA GEMA, Enquête avril 2014
- « L'essentiel du RSI en chiffres », RSI, données 2013
- « QIS5 : Calibration des risques en assurance de personne », Frédéric Planchet, septembre 2010

Mémoires d'actuariat

- « Construction de tables d'expérience pour l'entrée et le maintien en incapacité », Sophie Marton, CNAM 2007
- « Etude du risque arrêt de travail : construction de tables de maintien en incapacité, invalidité et probabilités de passage en invalidité », Séverine Marillet, ISFA 2004

Cours

- « Modèles de durée, Estimation du maintien en arrêt de travail », Frédéric Planchet, ISFA 2008