

**Mémoire présenté devant le CNAM pour l'obtention du Master
Droit Economie Gestion, mention Actuariat
et l'admission à l'Institut des Actuaires**

le 23 mars 2022

Par : Slim BELGHITH

Titre:

Provisionnement du maintien en incapacité : cas d'un portefeuille de
contrats caractérisés par une franchise courte

Confidentialité : NON OUI Durée : 2 ans

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

Président du Jury :

M. Nathanaël ABECERA

signatures

Entreprise :

Nom : AG2R La Mondiale

Directeur de mémoire en entreprise :

*Membres présents du jury de
l'Institut des Actuaires :*

Mme Sophie MICHON

M. Florent PERNOUD

Nom : Michael JACCAZ

Signature :

Invité :

Nom :

Signature :

*Membres présents du jury du
Cnam :*

M. Olivier DESMETTRE

**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

Secrétariat :

Bibliothèque :

Signature du candidat

Table des matières

Table des figures	6
Liste des tableaux	7
Liste des abréviations	8
Remerciements	9
Résumé	10
Abstract	11
Introduction	12
Partie 1 : Cadre du mémoire	13
Chapitre 1 : Présentation générale de la prévoyance	14
1. Organisation de la prévoyance	14
2. Les garanties couvertes	15
Chapitre 2 : Présentation du risque incapacité	16
1. Les prestations de la Sécurité Sociale	16
2. Les prestations complémentaires à la Sécurité Sociale	17
3. Présentation de la notion de franchise	17
Chapitre 3 : Principe de provisionnement du risque arrêt de travail	17
1. Cadre réglementaire	17
2. Description des tables réglementaires utilisées	19
3. Calcul des provisions mathématiques	20
Chapitre 4 : Exposition de la problématique à étudier	22
Partie 2 : Données utilisées	25
Chapitre 1 : Description de la base de données	26
Chapitre 2 : Fiabilisation et traitement des données	26
1. Traitement des doublons	27
2. Contrôles de cohérence entre variables	27
2.1 Fiabilisation de la variable « Franchise »	27
2.2 Contrôle de la cohérence entre la date de naissance de l'adhérent et le numéro de Sécurité Sociale	27
2.3 Contrôle de la cohérence entre le sexe de l'adhérent et le numéro de Sécurité Sociale ...	27
2.4 Contrôle de la cohérence de chronologie entre les dates	27
3. Analyse des valeurs manquantes	28
4. Détection des valeurs aberrantes	28

4.1 Contrôle de l'âge de l'assuré à la date de survenance	28
4.2 Contrôle de la durée de l'arrêt	28
Chapitre 3 : Analyse descriptive des données.....	29
1. Statistiques globales.....	29
2. Statistiques par année/mois de survenance	29
3. Statistiques par sexe.....	31
4. Statistiques par âge à la survenance de l'arrêt	32
5. Statistiques par franchise	33
6. Statistiques par type d'arrêt.....	34
7. Statistiques par CSP.....	35
8. Statistiques par situation familiale.....	36
9. Statistiques par salaire de référence.....	37
10. Statistiques par secteur d'activité	37
11. Analyse de la durée d'arrêt	38
12. Synthèse de l'analyse descriptive.....	39
Partie 3 : Construction de la loi de maintien en incapacité	40
Chapitre 1 : Sélection des variables discriminantes.....	41
Chapitre 2 : Choix de la période d'observation.....	43
Chapitre 3 : Présentation des notions de censure et de troncature	44
Chapitre 4 : Estimation des taux bruts de sortie d'incapacité.....	45
1. Présentation théorique de l'estimateur de Kaplan-Meier	45
2. Analyse des taux de sortie d'incapacité obtenus	48
Chapitre 5 : Lissage des taux de sortie d'incapacité	50
1. Présentation théorique de la méthode de Whittaker-Henderson.....	51
2. Description des tests d'adéquation de l'ajustement	54
3. Application de la méthode de Whittaker-Henderson aux données et validation de l'ajustement.....	55
4. Détermination de la loi de maintien en incapacité	59
Chapitre 6 : Prolongement de la loi d'expérience.....	59
Partie 4 : Mise à l'épreuve de la loi d'expérience.....	62
Chapitre 1 : Comparaison avec la table règlementaire	63
Chapitre 2 : Utilisation de la loi d'expérience	65
1. Calcul des provisions de maintien en incapacité.....	65
2. Etude d'impact sur le provisionnement	68
Chapitre 3 : Adaptation de la règle de sélection des encours	70
Conclusion.....	73

Bibliographie	74
Note de synthèse	75
Executive summary	78

Table des figures

Figure 1- Les différents états en arrêt de travail	18
Figure 2- Table de maintien en incapacité du BCAC	19
Figure 3- Table de passage en invalidité du BCAC.....	20
Figure 4- Table de maintien en invalidité du BCAC	20
Figure 5- Evolution du S/P incapacité par année de survenance et par date d'inventaire	23
Figure 6- Répartition des arrêts par année de survenance.....	30
Figure 7- Répartition des arrêts par mois de survenance	30
Figure 8- Durée moyenne des arrêts par année de survenance	31
Figure 9- Répartition des arrêts par âge et par sexe	32
Figure 10- Durée moyenne des arrêts par âge et par sexe	33
Figure 11- Répartition des arrêts par durée de franchise	33
Figure 12- Zoom sur la durée moyenne de l'arrêt en fonction de la franchise.....	34
Figure 13- Répartition des arrêts par type	34
Figure 14- Zoom sur la durée moyenne de l'arrêt par type	35
Figure 15- Répartition des arrêts par situation familiale	36
Figure 16- Zoom sur la durée moyenne de l'arrêt par situation familiale	37
Figure 17- Proportion des arrêts clos en fonction de l'ancienneté	38
Figure 18- Schématisation de la méthode des forêts aléatoires.....	41
Figure 19- Classement des variables discriminantes avec la méthode des forêts aléatoires	42
Figure 20- Répartition des arrêts par âge.....	43
Figure 21- Taux de sortie bruts sur 36 mois	48
Figure 22- Taille de l'intervalle de confiance à 95% des taux de sortie	49
Figure 23- Taux de sortie bruts sur 12 mois	49
Figure 24- Taux de sortie bruts sur 12 mois : zoom par âge	50
Figure 25- Taux de sortie lissés sur 12 mois	55
Figure 26- Comparaison des taux de sortie bruts et lissés.....	56
Figure 27- Taux lissés sur 12 mois : zoom par âge	57
Figure 28- Sorties modélisées vs. Sorties observées.....	58
Figure 29- Sorties modélisées vs. Sorties observées : zoom par âge	58
Figure 30- Extrait de la loi d'expérience	59
Figure 31- Illustration du prolongement	60
Figure 32- Loi de maintien en incapacité : zoom par âge.....	61
Figure 33- Comparaison des lois de maintien en incapacité.....	63
Figure 34- Comparaison des lois de maintien en incapacité : zoom par âge	63
Figure 35- Comparaison des taux de sortie pour un âge d'entrée en incapacité de 50 ans	64
Figure 36- Extrait du barème de provisionnement	65
Figure 37- Evolution de la provision unitaire en fonction de l'âge et de l'ancienneté dans l'arrêt.....	66
Figure 38- Evolution de la provision unitaire en fonction de l'ancienneté en arrêt : zoom par âge	66
Figure 39- Comparaison des barèmes de provisionnement	67
Figure 40- Illustration de la méthode de détermination du coefficient de provisionnement	68

Liste des tableaux

Tableau 1- Nombre d'encours et montant de PM incapacité par inventaire	22
Tableau 2- Analyse boni-mali du risque incapacité pour les inventaires 2018 et 2019.....	23
Tableau 3- Synthèse des traitements effectués sur la base de données initiale	29
Tableau 4- Principales statistiques descriptives du portefeuille	29
Tableau 5- Statistiques sur les arrêts par sexe	31
Tableau 6- Statistiques sur les arrêts par CSP	35
Tableau 7- Statistiques sur les arrêts par tranche de salaire	37
Tableau 8- Statistiques sur les arrêts par secteur d'activité	38
Tableau 9- Comparaison entre la survenance 2019 et la survenance 2018 à date d'arrêté identique	69
Tableau 10- Paramètres techniques pour le calcul des provisions	69
Tableau 11- PM incapacité en cours au 31/12/N.....	69
Tableau 12- Analyse boni/mali BCAC vs. Expérience	70
Tableau 13- Impact du changement de la règle de sélection des encours sur le calcul des PM au 31/12/2019.....	71
Tableau 14- PM incapacité en cours au 31/12/N (sélection des encours au 31/03/N+1)	71
Tableau 15- Analyse boni/mali BCAC vs. Expérience (sélection des encours au 31/03/N+1)	71

Liste des abréviations

ANI	Accord National Interprofessionnel
BCAC	Bureau Commun des Assurances Collectives
CART	Classification And Regression Trees
CCN	Convention Collective Nationale
CSP	Catégorie Socioprofessionnelle
CTIP	Centre Technique des Institutions de Prévoyance
PASS	Plafond Annuel de la Sécurité Sociale
PM	Provision Mathématique
SMIC	Salaire Minimum de Croissance
TME	Taux Moyen d'Emprunt d'Etat

Remerciements

En premier lieu, je tiens à exprimer mes remerciements envers toutes les personnes qui ont contribué de près ou de loin au bon déroulement de ce travail.

Je tiens particulièrement à remercier Michael JACCAZ, responsable du pôle prévoyance à la direction actuariat et acceptation médicale d'AG2R La Mondiale, pour sa disponibilité et ses remarques constructives tout au long de la réalisation de ce mémoire.

Mes remerciements s'adressent aussi à mes responsables à la direction de la souscription d'AG2R La Mondiale pour m'avoir fait confiance.

Enfin, j'adresse mes remerciements à l'ensemble de mes relecteurs.

Résumé

L'objectif de ce mémoire est de construire une loi de maintien en incapacité sur un portefeuille de contrats d'assurance caractérisés par une franchise courte. Sur ce périmètre, nous constatons que l'utilisation de la loi réglementaire conduit à surestimer les provisions mathématiques, notamment sur la dernière survenance.

Pour modéliser la loi d'expérience, nous avons opté pour des méthodes non paramétriques. L'estimateur de Kaplan-Meier a été ainsi retenu pour la construction des taux bruts de sortie d'incapacité. Ensuite, un lissage de ces taux par la méthode de Whittaker-Henderson a été effectué pour diminuer les irrégularités. Enfin, à cause de la volumétrie des données, des raccordements avec la table du BCAC ont été nécessaires.

Pour tester la fiabilité de notre modélisation, nous avons calculé les provisions mathématiques de maintien en incapacité et nous avons évalué l'impact que représente l'utilisation de la nouvelle loi plutôt que la loi du BCAC.

Mots clés : incapacité, provisionnement, franchise courte, loi d'expérience, Kaplan-Meier, Whittaker-Henderson.

Abstract

The regulatory table is usually used in temporary disability reserving. In case of short deductible, this table overestimate the technical reserves. The purpose of this study is to build an experience table suited to those contracts.

For modeling the experience table, we chosed non-parametric methods. The Kaplan-Meier estimator was retained for the construction of the crude discharge rates from temporary disability. Then, these rates were smoothed by the Whittaker-Henderson method to reduce the irregularities. The final results were extended with the BCAC's table.

To test the reliability of our model, we calculated the reserves for temporary disability in progress and we assessed the impact of using the new table rather than the BCAC's table.

Keywords: temporary disability, reserving, short deductible, experience table, Kaplan-Meier, Whittaker-Henderson.

Introduction

Pour pouvoir honorer leurs engagements vis-à-vis de leurs assurés, les assureurs doivent constituer des provisions représentatives des risques auxquels ils sont exposés.

En prévoyance, la constitution des provisions pour sinistres en cours constitue un exercice crucial pour les organismes assureurs. Ces provisions, constituées afin de faire face aux versements futurs des prestations en cours, font appel au calcul des probabilités ainsi qu'à l'actualisation des flux financiers.

Plus précisément en arrêt de travail, le provisionnement des sinistres en cours est généralement effectué à l'aide des tables réglementaires établies par le Bureau Commun des Assurances Collectives. Ces tables sont basées sur des portefeuilles des principales compagnies d'assurance françaises.

Cependant, si la population assurée présente un caractère atypique, l'utilisation de ces tables peut conduire à modéliser une situation non représentative du risque. Elle se traduit alors par un sous-provisionnement ou un sur-provisionnement : C'est le cas des contrats de prévoyance collective qui prévoient une garantie incapacité avec une franchise courte.

Dans ce mémoire, nous nous intéressons aux franchises continues dont le délai ne dépasse pas 10 jours. Il s'agit d'une particularité présente dans plusieurs contrats collectifs assurés par AG2R La Mondiale. Cette spécificité est liée à des obligations de certains régimes de branches.

Sur ce périmètre, nous constatons que l'utilisation des tables réglementaires conduit à surestimer les provisions mathématiques et notamment les provisions d'incapacité en cours sur la dernière survenance. Des bonis récurrents et significativement élevés impactent considérablement les comptes de résultats de ces régimes.

Dans cette optique s'inscrit la volonté de construire une table de maintien en incapacité qui reflète au mieux la sinistralité de ce portefeuille. L'objectif sera d'évaluer plus précisément les provisions correspondantes et par conséquent assurer un meilleur pilotage des régimes concernés.

Ce mémoire est organisé en quatre grandes parties :

- La première partie sert à introduire le cadre et le périmètre de l'étude. Elle expose, notamment, la problématique du sujet à travers un exemple pratique ;
- La deuxième partie est consacrée à une analyse exploratoire du portefeuille étudié. Elle vise à synthétiser l'information contenue dans nos données ;
- La troisième partie précise la démarche suivie dans la construction de la loi d'expérience. Cette démarche est appliquée en accordant une attention particulière à l'adéquation du modèle à nos données ;
- Enfin, une étude comparative avec la table règlementaire existante est réalisée dans la quatrième partie.

Partie 1 : Cadre du mémoire

Dans cette partie, nous commençons par donner une vision générale de la prévoyance. Puis, nous présentons le risque incapacité de travail et notamment la notion de franchise.

Nous détaillons ensuite le cadre réglementaire du calcul des provisions mathématiques en arrêt de travail.

Enfin, nous exposons le contexte et la problématique de notre étude.

Chapitre 1 : Présentation générale de la prévoyance

1. Organisation de la prévoyance

La prévoyance regroupe toutes les opérations d'assurance qui permettent de se prémunir contre les aléas de la vie liés à la personne. Selon la définition officielle, elle regroupe « les opérations ayant pour objet la prévention et la couverture du risque décès, des risques portant atteinte à l'intégrité physique de la personne ou liés à la maternité, des risques d'incapacité de travail ou d'invalidité ou du risque chômage »¹.

En France, la prévoyance se décline sous 2 niveaux : une couverture de base et une couverture complémentaire.

La couverture de base est gérée par la Sécurité Sociale. Son but est d'assurer et généraliser la protection des individus contre les conséquences de certains risques (maladie, accident) ou de certaines situations (vieillesse, charges familiales). Les assurés sociaux sont regroupés sous différents régimes en fonction de leur catégorie professionnelle :

- Le régime général : couvre essentiellement les travailleurs salariés ou assimilés du secteur privé (environ 80% de la population) ;
- Le régime agricole : couvre les exploitants et les salariés agricoles ;
- Le régime des indépendants² : couvre les artisans, commerçants, industriels et professions libérales ;
- Les régimes spéciaux : couvrent les salariés qui ne font pas partie du régime général, comme les employés de la Banque de France, les agents de la SNCF ou encore ceux de la RATP.

Le financement de cette couverture se fait obligatoirement par tous les salariés et tous les employeurs via le versement de cotisations qui sont calculées généralement en fonction du salaire.

Dans la plupart des cas, les prestations de la Sécurité Sociale ne permettent pas de remplacer la perte de revenus ou la hausse des dépenses des individus concernés. Du fait de cette insuffisance, des régimes complémentaires de prévoyance sont nés. En contrepartie de cotisations, ces régimes ajoutent un complément de revenus aux indemnités versées par la couverture de base.

Un régime complémentaire de prévoyance peut être souscrit par une entreprise pour l'ensemble ou une partie de ses salariés. On parle alors de prévoyance collective, qui peut avoir un caractère obligatoire ou facultatif. Par ailleurs, une personne peut souscrire individuellement à un régime complémentaire non professionnel. On parle alors de prévoyance individuelle.

La mise en place d'un régime de prévoyance collectif par l'entreprise est obligatoire dans le cadre de :

La Convention Collective Nationale des cadres de 1947 : Il s'agit de la seule obligation interprofessionnelle en matière de prévoyance³. Elle impose à l'employeur de cotiser à hauteur de

¹ Article 1 de la loi n° 89-1009 du 31 décembre 1989, dite loi EVIN.

² Depuis le 1^{er} janvier 2020, la protection sociale des indépendants est intégrée au régime générale de la Sécurité Sociale.

³ Sans considérer l'Accord National Interprofessionnel de 2013 qui a généralisé la complémentaire santé depuis le 1^{er} janvier 2016.

1,50% TA⁴ pour les salariés cadres et assimilés, affectés prioritairement au risque décès. En cas de défaillance, l'employeur doit verser une prestation égale à trois fois le plafond annuel de la Sécurité Sociale aux ayants droit du salarié décédé.

La loi de mensualisation : La loi de généralisation du 19 janvier 1978 modifiée par l'Accord National Interprofessionnel (ANI) du 11 janvier 2008 impose à l'employeur de maintenir un niveau minimum de salaire pour ses salariés en arrêt de travail. La durée de maintien du salaire varie selon l'ancienneté du salarié dans l'entreprise. La couverture se déclenche à compter du 8^{ème} jour d'arrêt en cas de maladie ou accident de la vie privée. Le point de départ de l'indemnisation est ramené au 1^{er} jour de l'arrêt en cas de maladie professionnelle ou accident de travail. Pour bénéficier de cette garantie, il faut avoir une ancienneté minimum de 1 an dans l'entreprise.

Il est à noter que ces obligations légales ne s'appliquent pas en présence de dispositions conventionnelles plus favorables.

Les conventions collectives de branches : Les entreprises qui adhèrent à une convention collective de branche sont tenues d'appliquer les dispositions obligatoires qui en découlent. Ainsi, un régime complémentaire de prévoyance peut être imposé dans ce cas. Les conventions collectives peuvent améliorer les obligations de maintien de salaire et créer des obligations en matière de couverture complémentaire.

Il existe plus de 300 conventions collectives au niveau national (CCN). Selon une étude menée en janvier 2018 par le CTIP (Centre Technique des Institutions de Prévoyance), 215 branches avaient signé un accord prévoyance.

2. Les garanties couvertes

Quatre familles de garanties sont regroupées sous le terme de prévoyance (dans le sens large) :

- Les garanties liées au décès : ces garanties visent à compenser la perte de revenus des bénéficiaires suite au décès du salarié. Elles peuvent prendre 3 formes :
 - o Un capital décès versé aux bénéficiaires ;
 - o Une rente de conjoint, temporaire ou viagère, servie au conjoint survivant ;
 - o Une rente éducation versée au bénéfice des enfants à charge ;
- Les garanties de remboursement de frais de santé : il s'agit de prestations en nature qui assurent le remboursement partiel ou intégral des soins médicaux et paramédicaux (consultations et visites, optique, dentaire, hospitalisation, pharmacie, frais d'analyse, ...) ;
- Les garanties liées à l'arrêt de travail : il s'agit de prestations en espèces qui permettent la compensation d'une perte de revenus suite à un arrêt de travail. On distingue deux états lorsqu'on fait référence à un arrêt de travail :
 - o L'incapacité (incapacité temporaire de travail) : l'indemnisation est sous forme d'indemnités journalières ;
 - o L'invalidité (incapacité permanente de travail) : elle peut faire suite ou non à une période d'incapacité. L'indemnisation est sous forme de rentes ;
- La garantie dépendance⁵ : elle peut prendre la forme d'une rente viagère et/ou d'un capital. Destinée aux personnes en situation de perte d'autonomie, cette garantie permet de financer

⁴ Tranche A (ou encore Tranche 1) : correspond aux rémunérations dans la limite du plafond mensuel de la Sécurité Sociale (3 428 € en 2020).

⁵ On parle d'état de dépendance (ou perte d'autonomie) lorsqu'une personne a besoin de faire appel à un tiers pour effectuer un ou plusieurs actes de la vie quotidienne.

une aide à domicile, payer un hébergement dans un établissement spécialisé ou encore d'aménager le lieu de résidence.

A présent, nous allons faire un zoom sur le risque incapacité qui fait l'objet de ce mémoire.

Chapitre 2 : Présentation du risque incapacité

1. Les prestations de la Sécurité Sociale

D'après l'article L.321-1 du code de la Sécurité Sociale « l'assurance maladie assure le versement d'indemnités journalières à l'assuré qui se trouve dans l'incapacité physique constatée par le médecin traitant [...] de continuer ou de reprendre le travail ».

La durée d'indemnisation maximale est de 3 ans. L'incapacité est transformée en invalidité au terme de cette période⁶. Un salarié peut sortir de l'état d'incapacité pour plusieurs raisons : reprise d'activité⁷, décès et passage en invalidité.

Les conditions d'ouverture des droits varient selon la durée de l'arrêt :

- Pour une durée inférieure à 6 mois : le salarié doit avoir travaillé au moins 150 heures au cours des 3 derniers mois précédant l'arrêt de travail ou avoir cotisé sur 1 015 fois le SMIC horaire dans les 6 mois civils précédant l'arrêt ;
- Pour une durée supérieure à 6 mois : le salarié doit avoir 12 mois d'immatriculation en tant qu'assuré social auprès de l'assurance maladie. De plus, il doit avoir travaillé au moins 600 heures au cours des 12 derniers mois précédant l'arrêt de travail ou avoir cotisé sur 2 030 fois le SMIC horaire.

Les règles d'indemnisation de la Sécurité Sociale diffèrent selon la nature de l'arrêt :

- En cas de maladie : les indemnités journalières sont versées après un délai de carence de 3 jours. Elles sont égales à 50% du salaire journalier de base (les 3 dernières payes échues divisées par 91,25 limité à 1,8 fois le SMIC) ;
- En cas d'accident de travail : les indemnités journalières sont versées dès le premier jour d'arrêt (pas de carence). Elles sont égales à 60% du salaire journalier de base (la dernière paye échue divisée par 30,25 dans la limite de 0,834% du PASS⁸) jusqu'au 28^{ème} jour, puis 80% à partir du 29^{ème} jour.

Il est à noter que les indemnités journalières sont majorées à partir du 31^{ème} jour d'arrêt continu si le salarié a au moins 3 enfants à charge. Elles sont alors égales à 66,66% du salaire journalier de base en cas de maladie.

⁶ Article L.341-3 du code de la Sécurité Sociale.

⁷ La reprise d'activité à temps partiel pour motif thérapeutique permet de cumuler les indemnités journalières avec une activité rémunérée.

⁸ Plafond Annuel de la Sécurité Sociale.

2. Les prestations complémentaires à la Sécurité Sociale

La garantie incapacité prévoit en général le versement d'indemnités journalières complémentaires à celles de la Sécurité Sociale au-delà d'une période de franchise. Le niveau de garantie s'exprime souvent en fonction du salaire de référence défini dans le contrat. La durée d'indemnisation maximale est de 3 ans à compter de la date de survenance de l'arrêt.

Le principe de non-enrichissement en assurance s'applique bien évidemment pour les régimes de prévoyance complémentaire. En d'autres termes, le cumul des prestations perçues par l'assuré (régime de base et régime complémentaire) ne doit pas dépasser son salaire avant l'entrée en arrêt de travail.

3. Présentation de la notion de franchise

En incapacité, la franchise est la période au terme de laquelle débute l'indemnisation. Elle s'exprime en nombre de jours et peut être variable d'un contrat à un autre. Il existe plusieurs types de franchise :

- La franchise continue : l'arrêt commence à être indemnisé après un nombre déterminé de jours d'absence continue. C'est la franchise la plus courante ;
- La franchise discontinue : l'arrêt commence à être indemnisé dès qu'un nombre cumulé de jours d'absence est comptabilisé sur une période donnée (par exemple une année) ;
- La franchise rétroactive : l'arrêt commence à être indemnisé après un nombre déterminé de jours d'absence continue. Les indemnités correspondant à la période de franchise sont versées rétroactivement à l'issue de cette période.

La franchise peut être différente en fonction de l'origine de l'arrêt de travail (maladie, accident de travail, hospitalisation, ...).

Dans le cadre de ce mémoire, nous nous intéressons aux franchises continues dont le délai ne dépasse pas 10 jours. Il s'agit d'une particularité que nous retrouvons dans certains contrats collectifs. Cette spécificité est liée aux obligations de certaines Conventions Collectives Nationales (CCN).

Chapitre 3 : Principe de provisionnement du risque arrêt de travail

1. Cadre règlementaire

Afin de respecter ses engagements futurs, l'organisme assureur doit évaluer et provisionner le mieux possible les risques auxquels il est confronté. En effet, dès l'ouverture d'un sinistre, l'assureur doit être en mesure de payer les prestations correspondantes jusqu'à leur terme.

L'arrêté du 28 mars 1996 a fixé les règles de provisionnement des garanties incapacité et invalidité. Il est cité dans le code des assurances, le code de la Sécurité Sociale et le code de la mutualité.

En ce qui concerne le risque arrêt de travail, trois types de provisions mathématiques sont requises :

- Provision pour incapacité en cours : correspond aux prestations liées au maintien de l'assuré en incapacité ;
- Provision pour invalidité en attente : correspond aux rentes d'invalidité susceptibles d'intervenir ultérieurement au titre des sinistres incapacité en cours ;
- Provision pour invalidité en cours : correspond aux prestations liées au maintien de l'assuré en invalidité.

La somme de la provision pour incapacité en cours et la provision pour invalidité en attente constitue la provision totale d'incapacité.

Le calcul de ces provisions repose sur deux éléments essentiels :

- Un taux d'actualisation (appelé aussi taux technique) : il permet de tenir compte des intérêts financiers à acquérir. Le taux technique est réglementé. En effet, il ne peut pas excéder 75% du taux moyen au cours des 24 derniers mois des emprunts de l'Etat français (TME), sans pouvoir dépasser 4,5% ;
- Des lois de maintien/de passage : utilisation des tables réglementaires du Bureau Commun des Assurances Collectives (BCAC) ou des lois d'expériences certifiées par un actuinaire indépendant et agréé à cet effet.

Le schéma suivant permet d'explicitier les liens entre les différents états qu'un assuré peut avoir et les provisions présentées dans cette section :

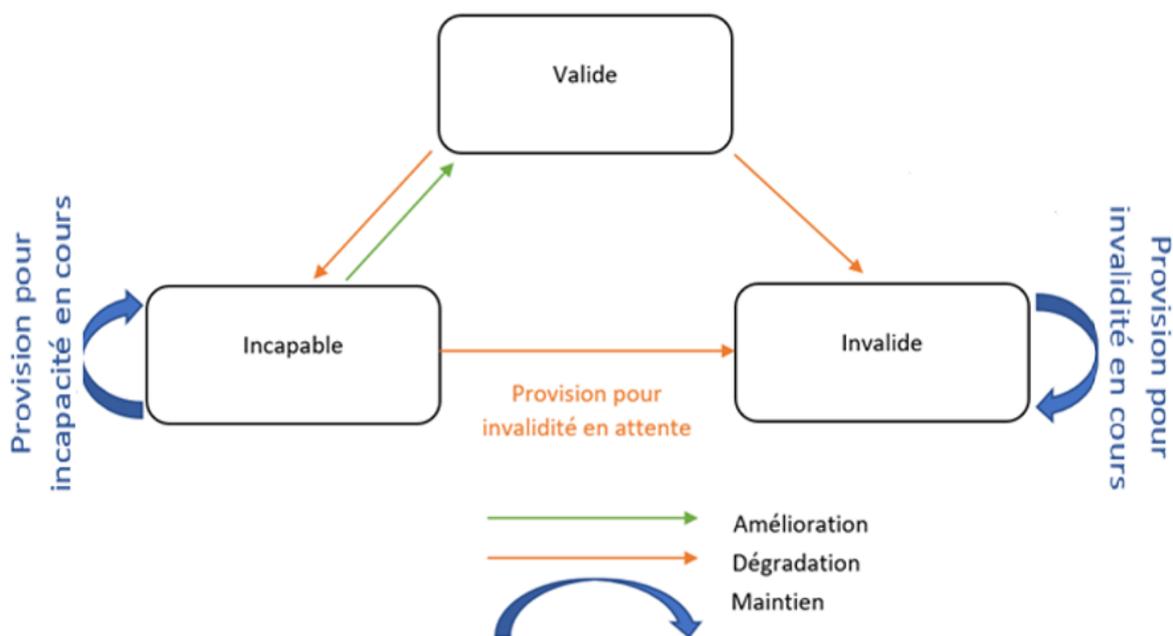


Figure 1- Les différents états en arrêt de travail

Remarques :

- La sortie de l'incapacité peut se faire par la reprise de travail (l'individu redevient valide), le passage en invalidité ou le décès ;
- L'état d'invalidité est irréversible contrairement à l'incapacité ;
- Une personne sort de l'invalidité lorsqu'elle part à la retraite ou décède.

2. Description des tables règlementaires utilisées

Les tables règlementaires ont été construites par le BCAC pour la première fois en 1993 à partir de données issues des portefeuilles des plus grandes compagnies d'assurance du moment. Une révision de ces tables a été nécessaire à la suite de la réforme des retraites⁹. Ainsi, l'arrêté du 24 décembre 2010 a permis de réviser les tables règlementaires en les prolongeant de deux ans. Pour prendre en compte l'évolution du risque arrêt de travail, le BCAC a construit de nouvelles tables en 2013. Toutefois, ces tables ne sont pas encore homologuées et par conséquent, elles ne peuvent pas remplacer les tables de 2010 qui restent toujours en vigueur.

Nous présentons ci-dessous les tables de maintien en incapacité, passage en invalidité et maintien en invalidité du BCAC (2010). Il est à noter que ces tables sont bidimensionnelles : elles tiennent compte de l'âge d'entrée en incapacité/invalidité et de l'ancienneté dans cet état. Toutefois, elles ne sont pas segmentées par sexe contrairement aux tables de mortalité.

Loi de maintien en incapacité :

En ligne : l'âge à l'entrée en incapacité.

En colonne : l'ancienneté en mois dans l'état d'incapacité.

L'âge varie de 23 ans ou moins à 66 ans. L'ancienneté est comprise entre 0 et 36 mois.

		Ancienneté dans l'état d'incapacité (en mois)					
		0	1	2	3	...	36
Age à l'entrée en incapacité	23 ou -	10000	2842	1743	1144	...	15
	24	10000	2931	1848	1215	...	14
	25	10000	3080	2001	1345	...	16
	26	10000	3177	2112	1461	...	23

	66	10000	5611	2427	1968	...	433

Figure 2- Table de maintien en incapacité du BCAC

Dans chaque élément de la table, figure le nombre de personnes encore en incapacité sur un effectif de 10 000 incapables présents au départ, pour l'âge à l'entrée et l'ancienneté considérée.

Exemple : Parmi les 10 000 individus entrés en incapacité à l'âge de 25 ans, il ne reste plus que 1 345 au bout de 3 mois.

Loi de passage en invalidité :

En ligne : l'âge à l'entrée en incapacité.

En colonne : l'ancienneté en mois dans l'état d'incapacité.

L'âge varie de 20 ans ou moins à 66 ans. L'ancienneté est comprise entre 0 et 35 mois.

⁹ La loi n°2010-1330 du 9 novembre 2010 portant réforme des retraites a reculé l'âge de départ à la retraite de deux ans.

		Ancienneté dans l'état d'incapacité (en mois)					
		0	1	2	3	...	35
Age à l'entrée en incapacité	20 ou -	1	0	0	0	...	39
	21	1	0	0	0	...	39
	22	1	0	0	0	...	39
	23	1	0	0	0	...	39

	61	15	13	4	7	...	305

Figure 3- Table de passage en invalidité du BCAC

Dans chaque élément de la table, figure le nombre de personnes passées en invalidité sur un effectif de 10 000 incapables présents au départ, pour l'âge à l'entrée et l'ancienneté considérée.

Exemple : Parmi les 10 000 individus entrés en incapacité à l'âge de 61 ans, 7 sont reconnus invalides par la Sécurité Sociale au cours du 3^{ème} mois.

Loi de maintien en invalidité :

En ligne : l'âge à l'entrée en invalidité.

En colonne : l'ancienneté en année dans l'état d'invalidité.

L'âge va de 20 ans à 61 ans. L'ancienneté est comprise entre 0 et 42 ans. La table est triangulaire supérieure gauche (âge + ancienneté <= 62 ans¹⁰).

		Ancienneté dans l'état d'invalidité (en année)					
		0	1	2	3	...	42
Age à l'entrée en invalidité	20	10000	9859	9699	9534	...	5663
	21	10000	9859	9699	9534	...	
	22	10000	9859	9699	9534	...	
	23	10000	9859	9699	9534	...	
	
	61	10000	9831				

Figure 4- Table de maintien en invalidité du BCAC

Dans chaque élément de la table, figure le nombre de personnes encore en invalidité sur un effectif de 10 000 invalides présents au départ, pour l'âge à l'entrée et l'ancienneté considérée.

Exemple : Parmi les 10 000 individus entrés en invalidité à l'âge de 20 ans, il ne reste plus que 5 663 au bout de 42 ans.

3. Calcul des provisions mathématiques

Tout sinistre dont la prestation est versée périodiquement pendant une durée aléatoire doit être provisionné jusqu'au terme de la garantie.

Les provisions mathématiques représentent la valeur actuelle probable des engagements, au niveau atteint, de l'organisme assureur vis-à-vis de l'assuré. Ces provisions sont calculées tête par tête.

¹⁰ Correspond à l'âge légal de départ à la retraite.

Nous donnons ci-dessous les formules de calcul de chacune des provisions mathématiques pour le versement d'un euro de rente. Pour évaluer ces quantités (appelées barèmes), le BCAC suppose que les paiements interviennent en milieu de période : cela revient à prendre la moyenne du terme d'avance et du terme échu.

Notons :

i : taux d'actualisation annuel.

$L_{x,k}^{INC}$: nombre d'individus entrés en incapacité à l'âge x et toujours dans cet état dans k mois (donné par la loi de maintien en incapacité).

$L_{x,k}^{INV}$: nombre d'individus entrés en invalidité à l'âge x et toujours dans cet état dans k années (donné par la loi de maintien en invalidité).

$S_{x,k}$: nombre d'individus entrés en incapacité à l'âge x et qui deviennent invalides au cours du $k^{\text{ème}}$ mois (donné par la loi de passage en invalidité).

➤ **Calcul de la provision mathématique d'incapacité en cours**

Pour un individu entré en incapacité à l'âge x et encore dans cet état t mois plus tard :

$$PM_{x,t}^{INC} = \frac{1}{2 L_{x,t}^{INC}} \sum_{k=t}^{35} \frac{L_{x,k}^{INC}}{(1+i)^{\frac{k-t}{12}}} + \frac{L_{x,k+1}^{INC}}{(1+i)^{\frac{k+1-t}{12}}}$$

➤ **Calcul de la provision mathématique d'invalidité en cours**

Pour un individu entré en invalidité à l'âge x et encore dans cet état t années plus tard :

$$PM_{x,t}^{INV} = \frac{1}{2 L_{x,t}^{INV}} \sum_{k=t}^{62-x-1} \frac{L_{x,k}^{INV}}{(1+i)^{k-t}} + \frac{L_{x,k+1}^{INV}}{(1+i)^{k+1-t}}$$

➤ **Calcul de la provision mathématique d'invalidité en attente**

Pour un individu entré en incapacité à l'âge x et encore dans cet état t mois plus tard :

$$PM_{x,t}^{INC_INV} = \frac{1}{2 L_{x,t}^{INC}} \sum_{k=t}^{35} \left(\frac{S_{x,k}}{(1+i)^{\frac{k-t}{12}}} + \frac{S_{x,k+1}}{(1+i)^{\frac{k+1-t}{12}}} \right) * PM_{x+\frac{k+6}{12},0}^{INV}$$

Avec, $PM_{x+\frac{k+6}{12},0}^{INV}$ déterminé par interpolation entre $PM_{[x+\frac{k+6}{12}],0}^{INV}$ et $PM_{[x+\frac{k+6}{12}+1],0}^{INV}$ ¹¹.

Remarque : En plus des trois provisions citées précédemment, l'assureur doit constituer une provision au titre du maintien de la garantie décès pour les assurés en arrêt de travail. La réglementation n'impose pas une méthode de calcul bien déterminée. Le calcul de cette provision n'est pas détaillé dans ce mémoire.

¹¹ $[x]$ désigne la partie entière de x .

Chapitre 4 : Exposition de la problématique à étudier

Notre étude porte sur un portefeuille de contrats en prévoyance collective assurés par AG2R La Mondiale et qui sont caractérisés par une franchise courte de la garantie incapacité (inférieure ou égale à 10 jours).

Dans le cadre de la sortie des comptes clients, le provisionnement du risque arrêt de travail est actuellement effectué à l'aide des tables réglementaires¹² établies par le BCAC (Bureau Commun des Assurances Collectives).

Sur le périmètre étudié, nous constatons que l'utilisation des tables réglementaires conduit à surestimer les provisions d'incapacité en cours et d'invalidité en attente, notamment sur la dernière survenance.

Dans le cadre de ce mémoire, nous nous intéressons uniquement au risque de maintien en incapacité.

Pour illustrer ce constat, une comparaison du nombre de dossiers provisionnés et du volume de provisions incapacité à différentes dates d'inventaires est réalisée. A titre d'exemple, nous allons présenter dans ce qui suit les résultats de cette comparaison pour un régime conventionnel donné représentant environ 40% du périmètre étudié en termes de provisions mathématiques.

La garantie incapacité de ce régime fonctionne de la manière suivante : en cas de maladie ou accident de la vie privée (y compris trajet), indemnisation à compter du 4^{ème} jour d'arrêt. En cas de maladie professionnelle ou accident de travail, indemnisation à compter du 1^{er} jour d'arrêt.

Survenance	Inventaire 31/12/2017		Inventaire 31/12/2018		Inventaire 31/12/2019	
	Nombre encours ¹³	PM Incapacité ¹⁴	Nombre encours	PM Incapacité	Nombre encours	PM Incapacité
2015	55	164 523 €				
2016	150	1 042 076 €	53	181 093 €		
2017	992	7 160 711 €	156	1 186 847 €	74	274 786 €
2018			1 048	7 786 202 €	179	1 461 192 €
2019					1 038	8 868 965 €
Total	1 197	8 367 309 €	1 257	9 154 143 €	1 291	10 604 943 €

Tableau 1- Nombre d'encours et montant de PM incapacité par inventaire

Nous constatons que le nombre de dossiers provisionnés au titre de la survenance N diminue fortement à l'inventaire N+1. Ce phénomène est observé d'une façon récurrente d'un inventaire à un autre comme on peut le constater dans le tableau ci-dessus.

Une analyse boni-mali de liquidation permet de démontrer le constat de sur-provisionnement en confrontant les provisions constituées aux prestations réellement payées :

¹² Tables réglementaires - Annexes 1.1, 1.2 et 1.3 de l'article 331-22 du code des assurances mises à jour par l'arrêté du 24 décembre 2010.

¹³ Tout arrêt dont la date de dernière indemnisation est inférieure ou égale à 4 mois par rapport à la date d'inventaire est considéré comme un encours.

¹⁴ Le taux d'actualisation retenu pour le calcul des provisions mathématiques est de 0,50% au 31/12/2017, 0,62% au 31/12/2018 et 0,365% au 31/12/2019.

Boni (+) / Mali (-) ¹⁵		
Exercice	Inventaire 31/12/2018	Inventaire 31/12/2019
2015	16 929 €	
2016	143 482 €	-49 758 €
2017	3 584 385 €	103 188 €
2018		3 505 722 €
Total	3 727 867 €	3 559 152 €

Tableau 2- Analyse boni-mali du risque incapacité pour les inventaires 2018 et 2019

Cette analyse met en exergue un boni significativement important portée par la dernière survenance ce qui impacte considérablement les résultats du régime.

L'application d'une franchise courte en incapacité se traduit par une surestimation des provisions mathématiques de la dernière survenance lors de l'inventaire N puis un fort dégonflement de ces provisions lors de l'inventaire suivant.

Lors de l'établissement des comptes clients, un ratio S/P¹⁶ largement déficitaire est remarqué systématiquement sur la dernière survenance présentée. Ce ratio ne traduit pas la réalité de la situation et par conséquent, il ne permet pas de tirer des conclusions sur un éventuel déséquilibre technique du risque incapacité et d'une façon plus générale du régime assuré.

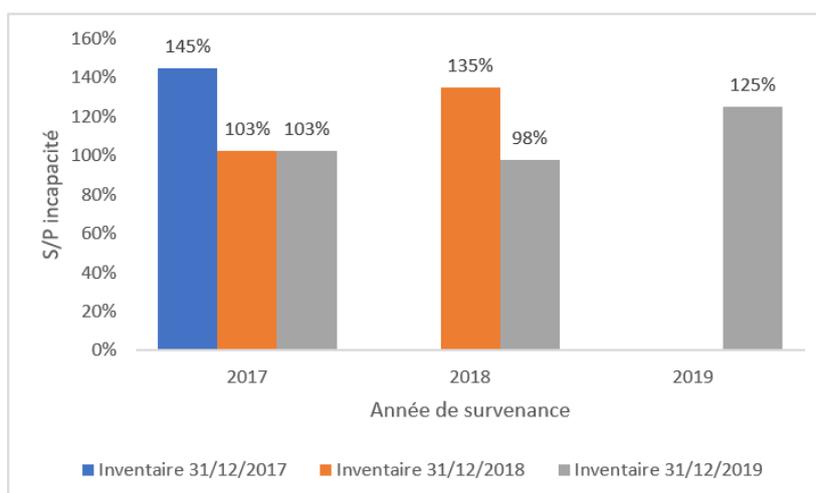


Figure 5- Evolution du S/P incapacité par année de survenance et par date d'inventaire

A travers cet exemple, nous avons mis en évidence que la table du BCAC est très prudente par rapport à la population assurée. C'est dans cette optique que s'inscrit la proposition de construire une loi d'expérience propre à notre portefeuille de franchise courte.

L'objectif sera donc de construire une loi de maintien en incapacité qui reflète d'une manière assez fidèle le comportement de la population assurée dans ce portefeuille.

L'utilisation de cette nouvelle loi permettrait de mieux évaluer les provisions liées à la garantie incapacité et par conséquent un meilleur pilotage des régimes concernés par la spécificité de la franchise courte.

¹⁵ Formule : (Prestations + PM) à l'inventaire N-1 – (Prestations + PM) à l'inventaire N.

¹⁶ Les ratios présentés n'intègrent pas de provision pour sinistres tardifs.

La construction d'une loi d'expérience fiable et robuste repose en grande partie sur le volume et la qualité des données à dispositions. La partie suivante leur est consacrée.

Partie 2 : Données utilisées

Dans cette partie, les données utilisées dans le cadre de l'étude seront détaillées. Une attention particulière sera accordée à la qualité de ces données. Les tests de fiabilisation ainsi que tous les traitements réalisés pour aboutir à la base finale seront précisés.

Après la fiabilisation des données, nous mènerons une analyse statistique profonde du portefeuille étudié. Cette analyse permettra de présenter quantitativement les variables à notre disposition tout en explicitant le lien avec le maintien en état d'incapacité de travail.

Chapitre 1 : Description de la base de données

La base d'étude initiale comporte 107 494 observations. Chaque observation représente un arrêt de travail survenu entre le 15/10/1998 et le 31/12/2018 dont la franchise est inférieure ou égale à 10 jours. Les données sont observées au 31/12/2019. Notre portefeuille est constitué de contrats collectifs en gestion directe issus de différentes Conventions Collectives Nationales (CCN).

Chaque observation est caractérisée par les variables suivantes :

- Numéro de contrat ;
- Identifiant de l'adhérent ;
- Date de survenance de l'arrêt ;
- Type de l'arrêt ;
- Date de première indemnisation ;
- Date de dernière indemnisation ;
- Numéro de Sécurité Sociale ;
- Nom de l'adhérent ;
- Prénom de l'adhérent ;
- Date de naissance de l'adhérent ;
- Sexe de l'adhérent ;
- Situation familiale de l'adhérent ;
- Catégorie socioprofessionnelle ;
- Salaire de référence ;
- Nom de la CCN.

Les variables suivantes sont créées pour la suite de l'étude :

- Durée de l'arrêt : différence (exprimée en jours) entre la date de dernière indemnisation et la date de survenance de l'arrêt ;
- Franchise : différence (exprimée en jours) entre la date de première indemnisation et la date de survenance de l'arrêt ;
- Age de l'assuré à la date de survenance : différence (exprimée en années) entre la date de survenance du sinistre et la date de naissance de l'assuré ;
- Etat du sinistre : Clos / En cours¹⁷.

Chapitre 2 : Fiabilisation et traitement des données

Le but de cette étape est de fiabiliser la base de données initiale pour aboutir à une base finale « propre » qui servira à la construction de notre loi d'expérience.

Une batterie de traitements permet de mener à bien cette phase :

¹⁷ Les sinistres dont la date de dernière indemnisation est antérieure au 01/09/2019 sont considérés comme clos en application de la règle de sélection des encours incapacité du groupe AG2R La Mondiale.

1. Traitement des doublons

Une clé constituée des variables « Numéro de Sécurité Sociale », « Date de survenance de l'arrêt », « Date de première indemnisation » et « Date de dernière indemnisation » permet de détecter les doublons présents dans la base. Nous avons procédé ainsi à la suppression de 1 678 observations. La présence d'un doublon peut être liée à l'une des raisons suivantes : changement de l'orthographe du nom ou du prénom de l'adhérent, transferts vers un nouveau contrat, modification du rattachement du contrat dans le périmètre de la branche, etc...

Par ailleurs, nous remarquons la présence de chevauchements des périodes indemnisées au titre de certains sinistres (pour un même adhérent et pour une même date de survenance). Pour ne pas biaiser les résultats, un regroupement sur une seule ligne a été réalisé en considérant la date de début d'indemnisation la plus ancienne et la date de dernière indemnisation la plus récente. Le nombre d'observations est ainsi réduit à 104 198.

2. Contrôles de cohérence entre variables

2.1 Fiabilisation de la variable « Franchise »

Un rapprochement entre la durée de la franchise calculée à partir de la base d'étude et celle de la franchise théorique (contractuelle) permet de mettre en évidence une adéquation à 99% entre ces deux valeurs. Par construction, notre base ne comporte que des arrêts avec des franchises limitées à 10 jours. Ainsi, les écarts sur les 1% observations restantes sont non significatifs.

2.2 Contrôle de la cohérence entre la date de naissance de l'adhérent et le numéro de Sécurité Sociale

Des anomalies ont été détectées à l'issue de ce test pour 30 observations. Etant donnée que nous ne pouvons pas prioriser aucune de ces deux variables et que le nombre d'anomalies est très faible, nous procédons à la suppression de ces lignes.

2.3 Contrôle de la cohérence entre le sexe de l'adhérent et le numéro de Sécurité Sociale

Des anomalies ont été détectées à l'issue de ce test pour 1 332 observations relatives à 880 individus. En examinant ces cas, nous constatons que le numéro de Sécurité Sociale est plus cohérent vis-à-vis du prénom de l'adhérent que son sexe. L'information apportée par le numéro de Sécurité Sociale sera donc priorisée pour ces 880 cas.

2.4 Contrôle de la cohérence de chronologie entre les dates

Ce test consiste à vérifier pour chaque ligne de la base si :

- La date de survenance de l'arrêt > la date de naissance de l'assuré ;
- La date de survenance de l'arrêt <= la date de première indemnisation ;
- La date de première indemnisation <= la date de dernière indemnisation.

Les 3 conditions sont respectées pour l'ensemble des observations.

3. Analyse des valeurs manquantes

Un test de détection des valeurs manquantes est réalisé pour les 19 variables de notre base. 30 valeurs manquantes pour la variable « Date de première indemnisation » ont été détectées. L'utilisation de la franchise théorique par contrat permet de compléter cette donnée à partir de la date de survenance. Toutes les autres variables sont renseignées à 100%.

4. Détection des valeurs aberrantes

4.1 Contrôle de l'âge de l'assuré à la date de survenance

Nous remarquons la présence de 119 observations relatives à des personnes ayant un âge inférieur à 18 ans. Il s'agit principalement des salariés en contrat d'apprentissage. L'âge minimum pour ce type de contrat est de 16 ans (voir 15 ans sous certaines conditions)¹⁸. 5 dates de naissance aberrantes (âge < 16 ans) sont toutefois détectées. Ces dates sont remplacées en se basant sur le numéro de Sécurité Sociale (le jour de naissance est fixé au 1^{er} du mois dans ce cas).

Par ailleurs, la base comporte 131 observations relatives à des personnes ayant un âge supérieur à 67 ans. Il s'agit des salariés qui n'ont pas encore liquidé leurs pension vieillesse de la Sécurité Sociale ou qui sont en situation de cumul emploi-retraite. En règle générale, l'assureur a l'obligation de maintenir la garantie du contrat sans aucune limite d'âge¹⁹. Ces observations sont donc conservées dans le périmètre de l'étude.

Enfin, 3 enregistrements comportent des dates de naissances non raisonnables au regard de la durée de vie humaine. Ces dates sont également remplacées en se basant sur le numéro de Sécurité Sociale.

4.2 Contrôle de la durée de l'arrêt

La base d'étude comporte 78 lignes relatives à des arrêts dont la durée dépasse 1 095 jours. Ces arrêts peuvent être liés à des rechutes : En effet, suite à une reprise d'activité consécutive à un arrêt, un assuré peut être amené à cesser de nouveau son activité pour la même cause que l'arrêt initial et dans un laps de temps déterminé. La période de franchise ne s'applique pas dans ce cas. Il s'agit d'un cas très rare surtout pour le portefeuille des contrats à franchise courte. Nous avons donc procédé à la suppression de ces lignes.

A la fin de cette étape, notre base d'étude finale est constituée. Elle contient 104 090 observations (soit 97% de la base initiale) sans doublons et sans valeurs manquantes. Un contrôle renforcé a permis de fiabiliser les données sensibles au calcul de l'âge à la survenance et la durée de l'arrêt. Une attention particulière a été ainsi accordée à la validité et la cohérence des variables de type date (naissance, survenance, première indemnisation, dernière indemnisation).

¹⁸ Articles L6222-1 à L6222-3 du code de travail.

¹⁹ Article L311-2 du code de la Sécurité Sociale.

Traitement	Nombre d'observations	Pourcentage
Suppression des doublons	1 678	1,56%
Regroupement des arrêts sur une seule ligne	1 618	1,51%
Suppression des arrêts dont la durée dépasse 1095 jours	78	0,07%
Suppression des observations présentant des incohérences	30	0,03%
Total suppressions	3 404	3,17%
Remplacement des valeurs manquantes	30	0,03%
Remplacement des dates de naissance aberrantes	8	0,01%
Correction de la variable sexe	1 332	1,24%
Total remplacements	1 370	1,28%

Tableau 3- Synthèse des traitements effectués sur la base de données initiale

Une fois la base d'étude fiabilisée, nous passons à l'analyse univariée de l'ensemble des variables disponibles.

Chapitre 3 : Analyse descriptive des données

1. Statistiques globales

Le tableau ci-dessous présente les principales statistiques descriptives du portefeuille :

Etat du sinistre	Nombre d'arrêts	Poids (en %)	Age moyen à l'arrêt	Durée moyenne de l'arrêt (en jours)
Clos	103 385	99%	41	68
En cours	705	1%	49	611
Total	104 090	100%	41	72

Tableau 4- Principales statistiques descriptives du portefeuille

Le portefeuille comporte 104 090 observations, dont 103 385 dossiers clos²⁰. L'âge moyen d'entrée en incapacité est de 41 ans. En moyenne, la durée d'arrêt est de 68 jours pour les dossiers clos.

2. Statistiques par année/mois de survenance

La répartition des arrêts par année de survenance est donnée dans le graphique qui suit :

²⁰ Pour rappel, les sinistres sont observés au 31/12/2019 avec une date de survenance limitée au 31/12/2018.

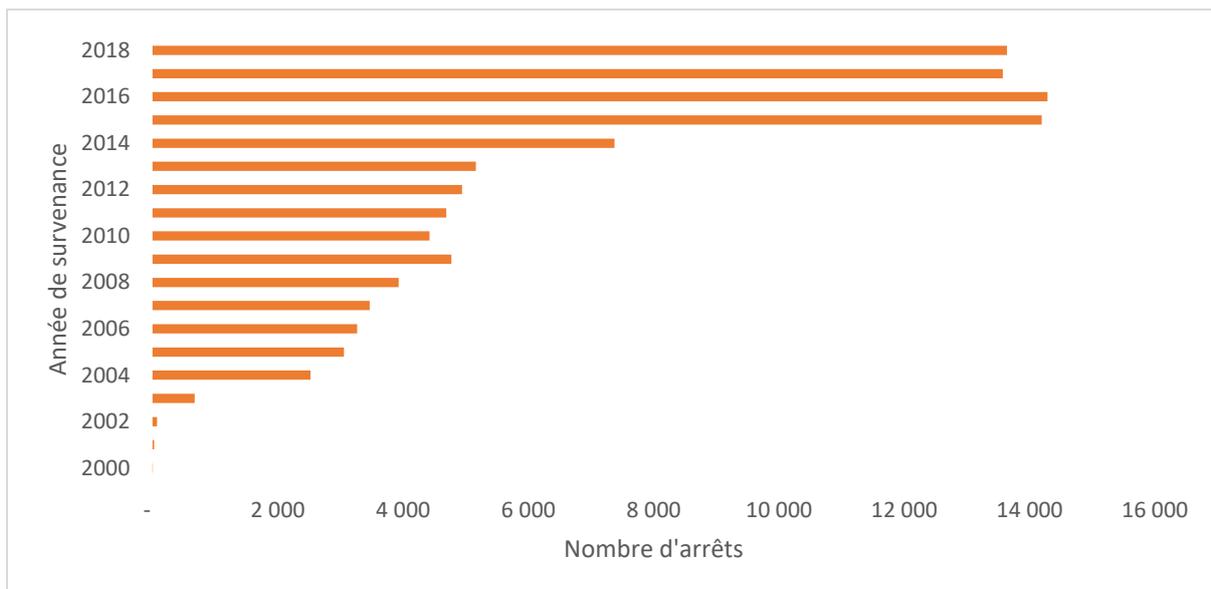


Figure 6- Répartition des arrêts par année de survenance

Nous constatons que le nombre d'arrêts de travail augmente globalement d'une année sur l'autre. Cette tendance s'explique par la croissance du nombre d'affaires nouvelles dans le portefeuille entre 2000 et 2016.

L'évolution importante du nombre d'arrêts entre 2014 et 2015 s'explique par la migration des contrats ex-PREMALLIANCE dans le système d'information d'AG2R La Mondiale en 2015.

Le graphique suivant présente la répartition des arrêts par mois de survenance :

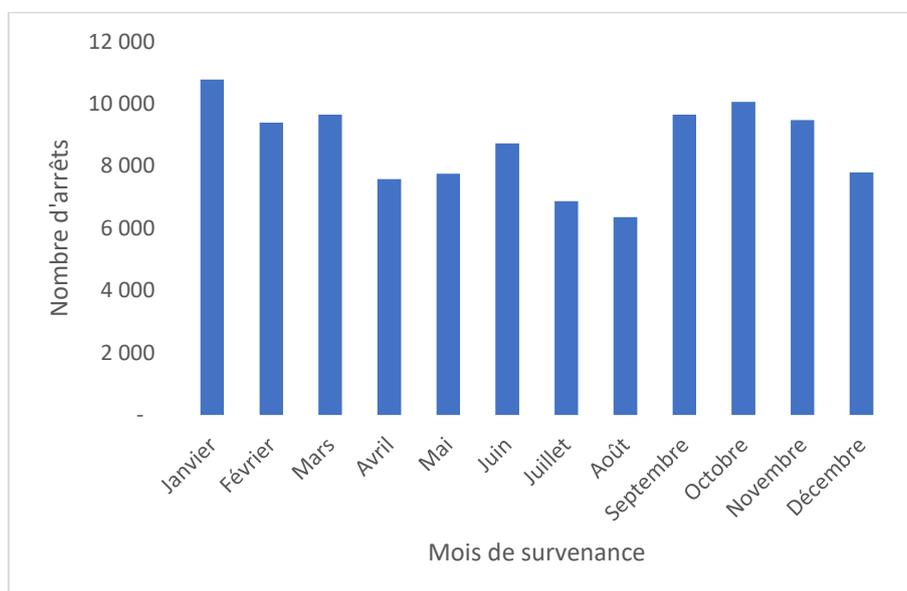


Figure 7- Répartition des arrêts par mois de survenance

Un effet saisonnalité des arrêts de travail est mis en exergue. Un pic de sinistralité est observé en début d'année (surtout au mois de janvier). A l'inverse, une chute de la sinistralité est observée sur les mois de juillet et août avant une reprise pendant l'automne. Des éléments climatiques ou encore psychologiques pourraient expliquer cette saisonnalité.

Nous représentons ci-dessous la durée moyenne de l'arrêt en fonction de l'année de survenance du sinistre (uniquement pour les dossiers clos) :

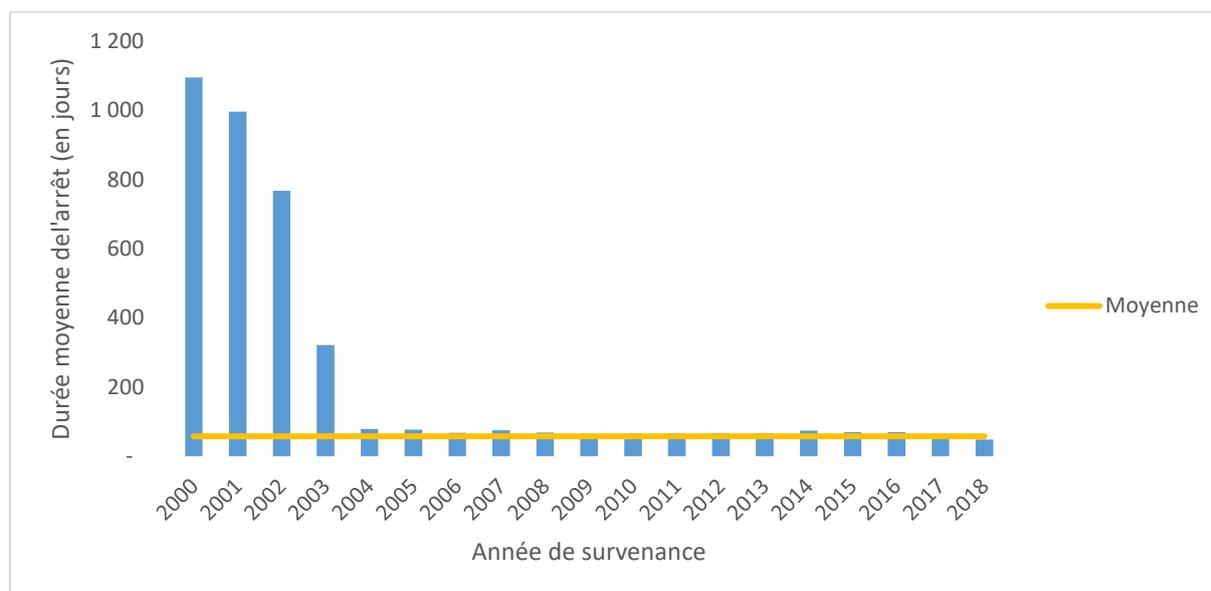


Figure 8- Durée moyenne des arrêts par année de survenance

Les durées moyennes observées entre 2000 et 2003 s'expliquent par un nombre très faible d'observations sur ces survenances (775 observations au total). A compter de l'année 2004, nous notons une stabilisation de la durée des arrêts par survenance avec des variations modérées autour d'une moyenne de 68 jours.

Le même constat a été fait pour l'évolution de l'âge moyen à l'arrêt en fonction de l'année de survenance qui n'est pas représentée ici.

Remarque : Pour ne pas biaiser notre analyse, nous ne prendrons en compte que les sinistres survenus après 2003 pour la modélisation du maintien en incapacité.

3. Statistiques par sexe

Les principales statistiques descriptives par sexe sont exposées dans le tableau ci-après :

Sexe	Nombre d'arrêts	Poids (en %)	Age moyen à l'arrêt	Durée moyenne de l'arrêt (en jours) *
Femme	92 064	88%	41	68
Homme	12 026	12%	42	67
Total	104 090	100%	41	68

Tableau 5- Statistiques sur les arrêts par sexe

*Durée calculée sur les dossiers clos

Les femmes sont largement majoritaires au sein de notre portefeuille. Elles représentent 88% du total des observations.

Nous notons que la durée moyenne de rester en état d'incapacité ainsi que l'âge moyen à l'arrêt sont sensiblement proches pour les deux sexes.

4. Statistiques par âge à la survenance de l'arrêt

La répartition des arrêts par âge à la survenance et par sexe est affichée dans le graphique suivant :

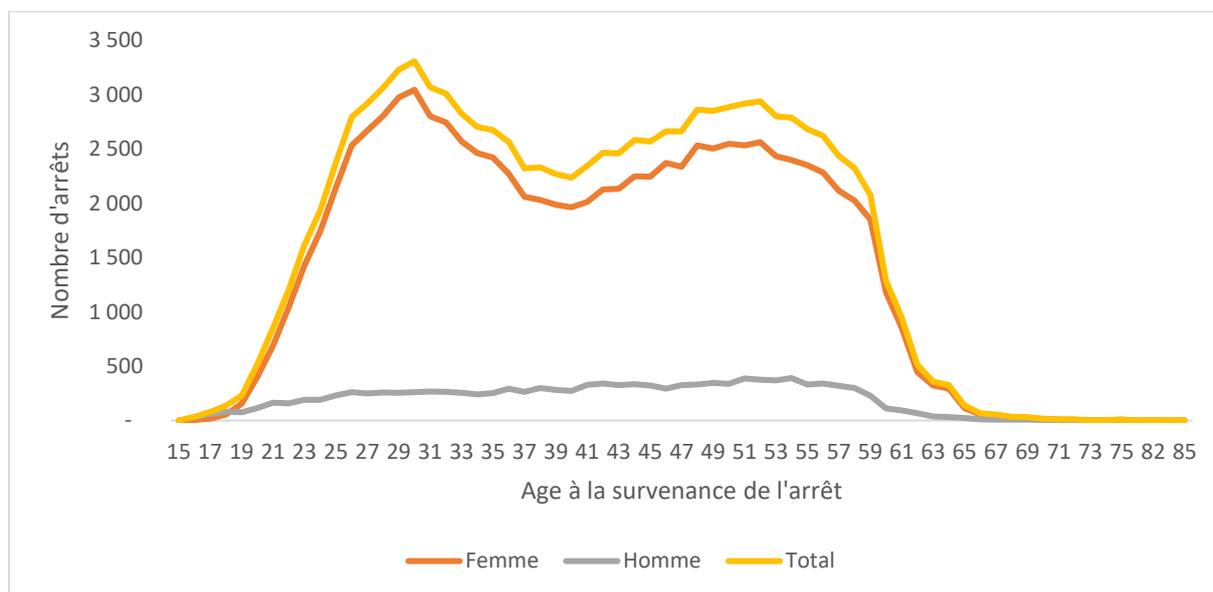


Figure 9- Répartition des arrêts par âge et par sexe

Nous observons un premier pic pour les arrêts chez les femmes à l'âge de 30 ans puis une diminution du nombre d'arrêts jusqu'à 40 ans pour remonter ensuite avec un deuxième pic à l'âge de 52 ans. Les arrêts chez les hommes se répartissent en fonction de l'âge d'une manière plus uniforme que chez les femmes.

Plus de 99% des arrêts sont concentrés entre 20 ans et 65 ans. Une décroissance forte du nombre d'arrêts est observée à partir de 60 ans ce qui pourrait être expliqué par les départs à la retraite.

Ci-après une présentation de la durée moyenne de l'arrêt en fonction de l'âge à la survenance et du sexe (uniquement pour les dossiers clos) :

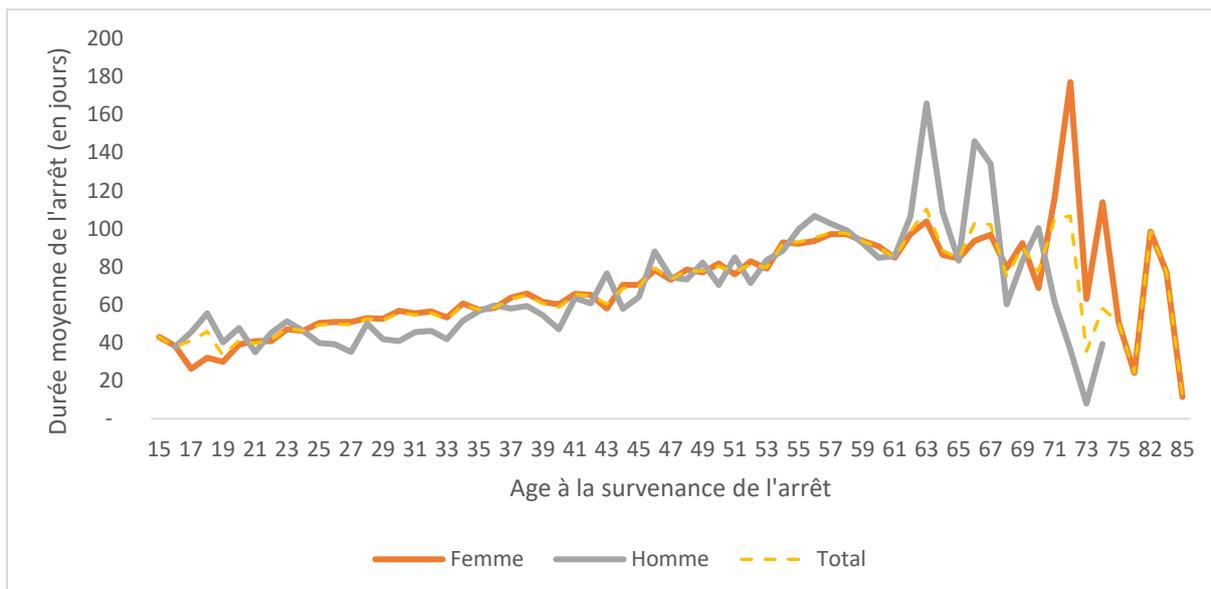


Figure 10- Durée moyenne des arrêts par âge et par sexe

L'analyse de la courbe de l'évolution de la durée moyenne de l'arrêt en fonction de l'âge montre une tendance haussière. La courbe des femmes se situe au-dessus de la courbe des hommes entre 25 ans et 40 ans. Nous pouvons donc conclure que la durée moyenne d'un arrêt est légèrement plus élevée chez les femmes que chez les hommes sur cette tranche d'âges.

Des irrégularités sont remarquées pour les âges les moins élevés (inférieurs à 20 ans) et pour les âges les plus élevés (supérieurs à 60 ans) en ce qui concerne la population masculine essentiellement. Ce constat est dû au faible nombre d'observations sur ces tranches d'âges.

5. Statistiques par franchise

Le graphique suivant détaille la répartition des arrêts par durée de franchise :

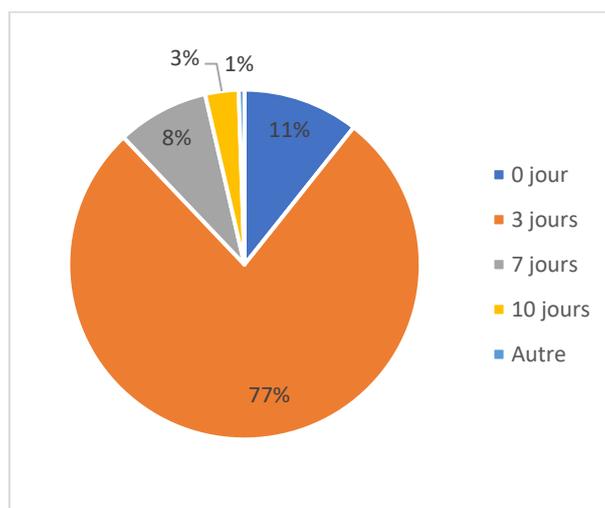


Figure 11- Répartition des arrêts par durée de franchise

Le regroupement "Autre" englobe les franchises intermédiaires entre 0 jour et 10 jours non représentées dans le graphique.

Nous remarquons que la proportion des arrêts d'une franchise de 3 jours est prépondérante dans la base d'étude (77%).

Nous affichons ci-après la durée moyenne de l'arrêt en fonction de la franchise (uniquement pour les dossiers clos) :

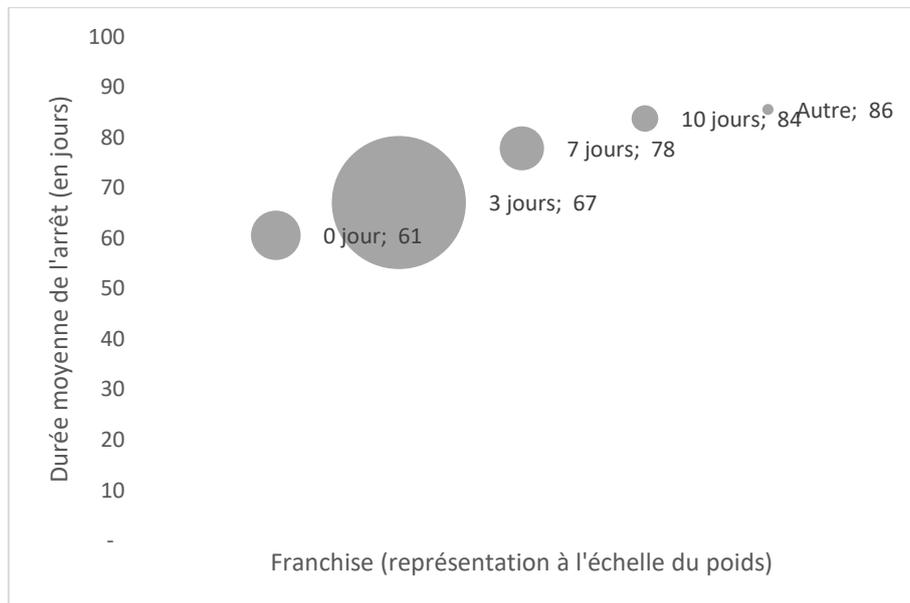


Figure 12- Zoom sur la durée moyenne de l'arrêt en fonction de la franchise

La variable « Franchise » semble être discriminante pour le maintien en incapacité. En effet, le graphique montre que la durée moyenne de l'arrêt croît avec la durée de la franchise.

6. Statistiques par type d'arrêt

La répartition des arrêts par type est donnée dans le graphique ci-après :

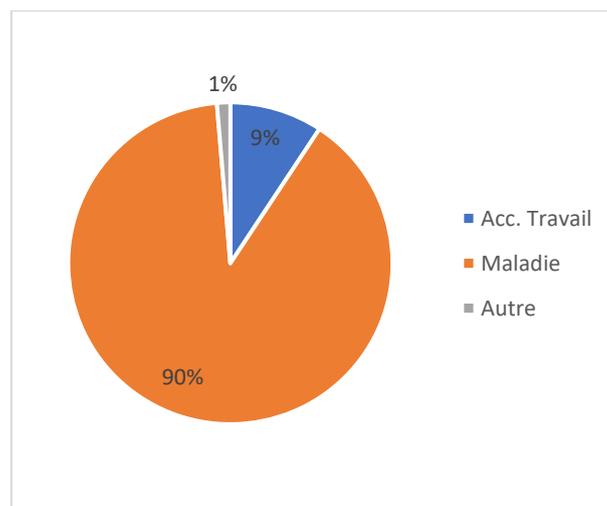


Figure 13- Répartition des arrêts par type

Le regroupement "Autre" englobe les types d'arrêts suivants : Accident de trajet, Accident de la vie privée, Affectation de longue durée, Longue maladie, Maladie hospitalière, Maladie professionnelle et Maternité.

La maladie représente la principale cause d'entrée en incapacité (90%) suivie par l'accident de travail (9%).

Remarque : Les arrêts de travail au titre des congés maternité (305 observations) seront exclues dans l'étape de modélisation. En effet, ces arrêts ont une durée indemnisable plafonnée selon la situation de l'assurée. Par ailleurs, ils font l'objet d'un provisionnement particulier selon les règles de calcul des provisions mathématiques adoptées par le groupe AG2R La Mondiale²¹.

Nous exposons ci-dessous la durée moyenne de l'arrêt en fonction du motif du sinistre (uniquement pour les dossiers clos) :

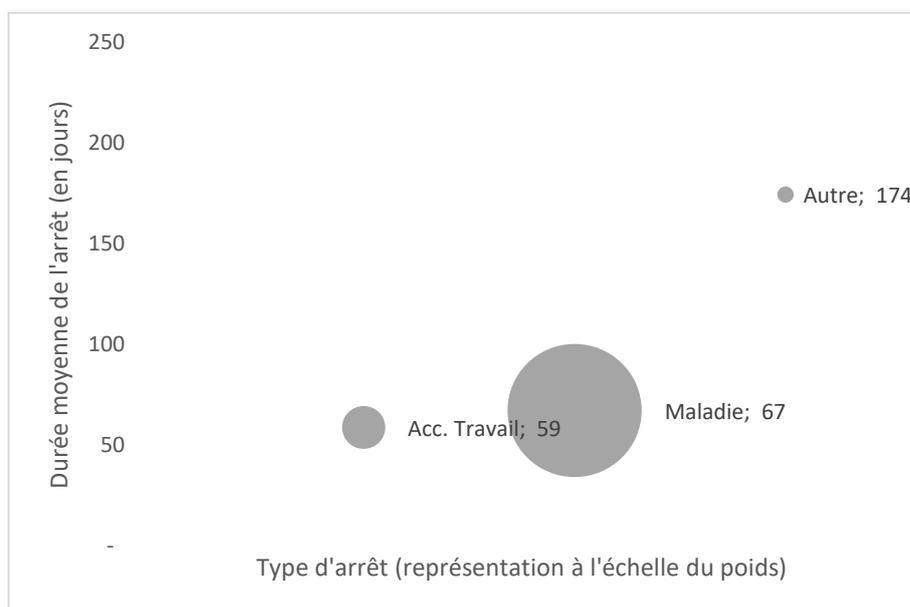


Figure 14- Zoom sur la durée moyenne de l'arrêt par type

Un arrêt engendré par une maladie dure en moyenne 8 jours de plus qu'un arrêt engendré par un accident de travail. Les sinistres de type « Autre » ont une durée moyenne plus élevée ce qui est cohérent avec les motifs de ces arrêts (hospitalisation, affectation de longue maladie, ...).

7. Statistiques par CSP

Le tableau ci-dessous présente les principales statistiques descriptives par CSP :

CSP	Nombre d'arrêts	Poids (en %)	Age moyen à l'arrêt	Durée moyenne de l'arrêt (en jours) *
Cadres	5 700	5%	38	67
Non cadres	98 390	95%	41	68
Total	104 090	100%	41	68

Tableau 6- Statistiques sur les arrêts par CSP

²¹ On considère que les congés maternité durent 196 jours et sont payés de manière certaine.

*Durée calculée sur les dossiers clos

Les arrêts concernent principalement les non cadres avec une proportion de 95% de la base des sinistrés. La durée moyenne de l'arrêt est stable par catégorie socioprofessionnelle. L'âge moyen à l'arrêt est légèrement moins élevé chez les cadres.

8. Statistiques par situation familiale

Le graphique suivant présente la répartition des arrêts par situation familiale :

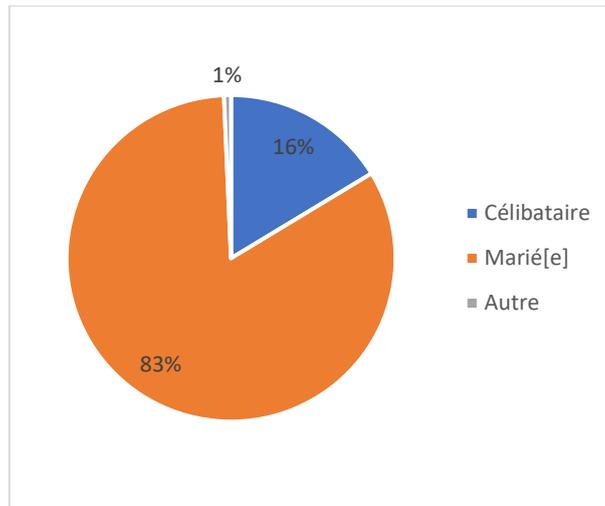


Figure 15- Répartition des arrêts par situation familiale

Le regroupement "Autre" englobe les situations suivantes : Divorcé(e), Veuf(ve) et Union libre.

Les mariés représentent 83% du portefeuille suivis par les célibataires avec 16%. Ce résultat est cohérent avec l'âge moyen de la population qui est de 41 ans.

Nous représentons ci-dessous la durée moyenne de l'arrêt en fonction de la situation familiale de l'assuré (uniquement pour les dossiers clos) :

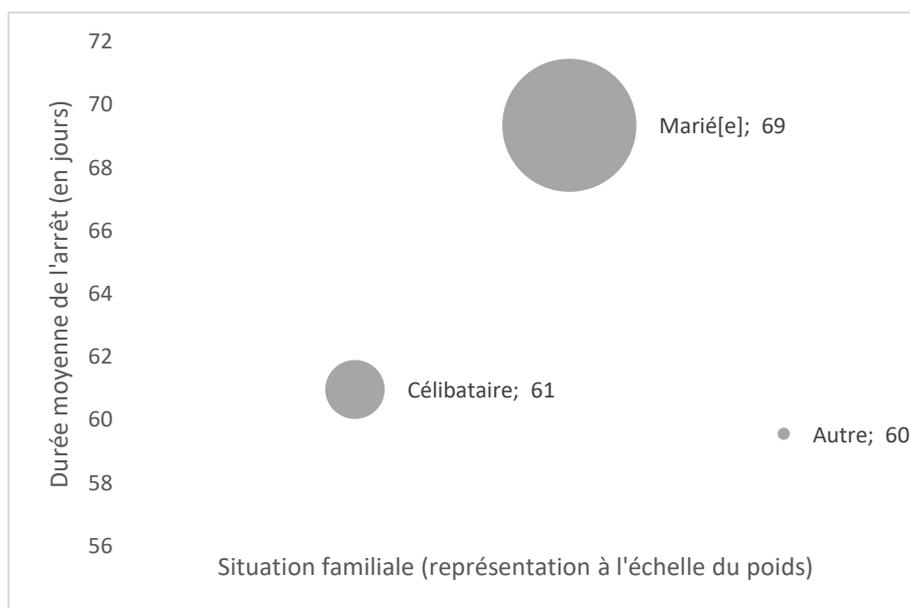


Figure 16- Zoom sur la durée moyenne de l'arrêt par situation familiale

Nous notons que les célibataires restent en moyenne 8 jours de moins que mariés en état d'incapacité. Ce résultat est attendu puisque les célibataires sont en moyenne plus jeunes (l'âge moyen des célibataires à l'arrêt est de 34 ans contre 43 ans pour les mariés).

9. Statistiques par salaire de référence

Le tableau ci-après met en avant les principales statistiques descriptives par tranche de salaire :

Tranche de salaire	Nombre d'arrêts	Poids (en %)	Age moyen à l'arrêt	Durée moyenne de l'arrêt (en jours) *
INF 17 K€	25 825	25%	42	81
17K€ - 21 K€	27 051	26%	39	66
21K€ - 26K€	27 396	26%	41	63
SUP 26 K€	23 818	23%	43	61
Total	104 090	100%	41	68

Tableau 7- Statistiques sur les arrêts par tranche de salaire

*Durée calculée sur les dossiers clos

Le salaire de référence moyen est de 22 K€. La médiane, qui est équivalente au salaire séparant en deux parties égale la population, s'établit à 21 K€.

A partir de ce tableau, nous pouvons remarquer que le salaire est inversement proportionnel à la durée moyenne de l'arrêt. Ce phénomène peut s'expliquer par la facilité d'accès aux soins des assurés les plus aisés financièrement ce qui permet une guérison plus rapide.

10. Statistiques par secteur d'activité

Le tableau qui suit présente les principales statistiques descriptives par secteur d'activité.

La variable « secteur d'activité » a été construite en se basant sur le nom des CCN (Conventions Collectives Nationales) présentes dans le périmètre étudié.

Secteur d'activité	Nombre d'arrêts	Poids (en %)	Age moyen à l'arrêt	Durée moyenne de l'arrêt (en jours) *
Santé	88 899	85%	41	66
Alimentaire	15 191	15%	42	77
Total	104 090	100%	41	68

Tableau 8- Statistiques sur les arrêts par secteur d'activité

*Durée calculée sur les dossiers clos

Le secteur de la santé est largement prépondérant au sein de notre portefeuille. Il représente 85% du total des observations contre 15% pour le secteur alimentaire.

Nous notons que la durée moyenne de rester en état d'incapacité dépend sensiblement du secteur d'activité malgré un âge moyen pratiquement constant. Quant au salaire de référence moyen, il est plus élevé dans le secteur de la santé (22 K€) que dans le secteur alimentaire (20 K€).

11. Analyse de la durée d'arrêt

Dans le graphique qui suit, nous représentons les taux de sorties cumulés de l'état d'incapacité par mois d'ancienneté pour l'ensemble du portefeuille.

Pour des raisons de lisibilité, seuls les résultats des 12 premiers mois sont affichés.

La convention suivante est utilisée : Nombre de jours par mois = 365,25/12.

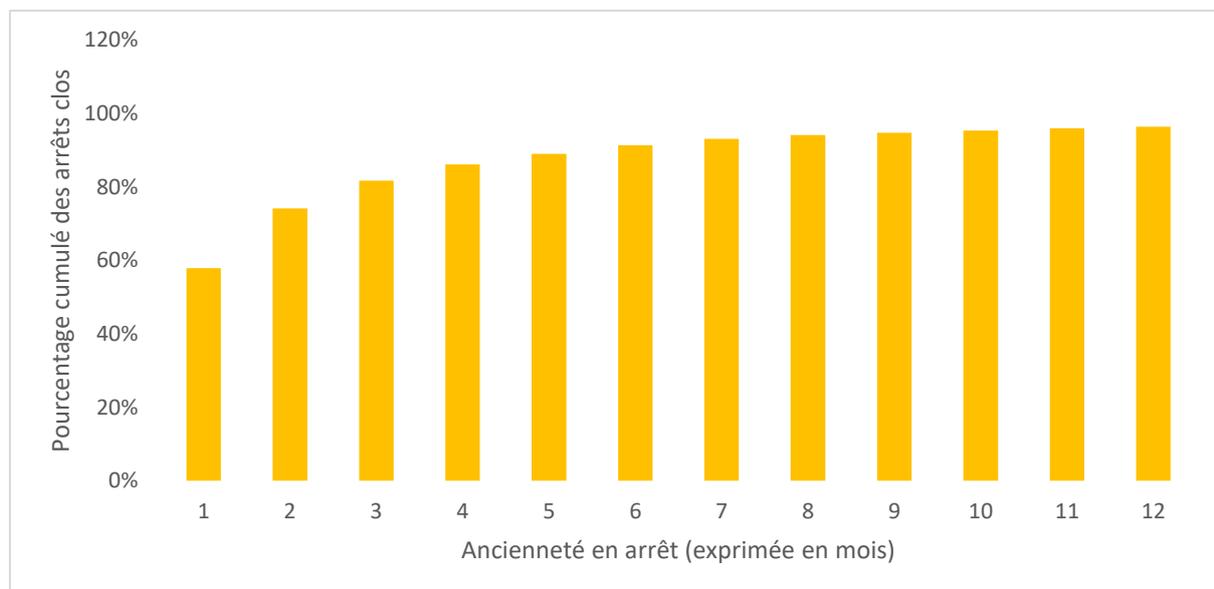


Figure 17- Proportion des arrêts clos en fonction de l'ancienneté

Les assurés dans notre portefeuille ont tendance à reprendre le travail assez rapidement, en moyenne après 68 jours. Les probabilités de sortie de l'état d'incapacité diminuent globalement en fonction de l'ancienneté de l'arrêt. Nous notons que :

- 58% des sinistres sont clos au bout d'un mois,
- 82% des sinistres sont clos au bout de 3 mois,
- 92% des sinistres sont clos au bout de 6 mois,
- 97% des sinistres sont clos au bout de 12 mois.

Les données sont très dispersées pour les anciennetés les plus élevées. Au-delà d'un certain seuil (à déterminer), un raccordement avec une table de référence nous paraît donc indispensable pour avoir une loi de maintien en incapacité sur 36 mois.

12. Synthèse de l'analyse descriptive

Les statistiques descriptives présentées dans cette partie nous ont permis de synthétiser l'information contenue au sein des données. Cette analyse exploratoire nous a permis d'avoir une vision du portefeuille dans sa globalité.

Il en ressort que la population étudiée est majoritairement féminine. Elle est constituée essentiellement de non cadres, mariés, exerçant dans le secteur de la santé. La principale cause d'entrée en état d'incapacité est la maladie. Nous avons noté aussi que l'âge d'entrée en incapacité est en moyenne autour de 41 ans et que 97% des sinistres sont clos dès la première année.

Les résultats statistiques ont également montré une certaine hétérogénéité de la population : La durée de maintien en incapacité semble dépendre de plusieurs facteurs comme l'âge à la survenance du sinistre, le salaire de référence de l'assuré, la durée de la franchise, ...

Une prise en compte de l'hétérogénéité est donc nécessaire pour disposer d'une loi d'expérience qui reflète au mieux la sinistralité du portefeuille étudié.

Partie 3 : Construction de la loi de maintien en incapacité

Dans ce qui suit, nous allons d'abord appliquer la méthode des forêts aléatoires pour choisir les variables les plus discriminantes du risque de maintien en incapacité.

Ensuite, nous allons procéder à la modélisation des taux bruts de sortie d'incapacité à l'aide de l'estimateur de Kaplan-Meier.

Pour répondre à la problématique de lissage de ces taux, la technique de Whittaker-Henderson sera utilisée.

Enfin, en raison de la volumétrie des données pour les anciennetés les plus élevées, nous allons proposer un prolongement avec une table de référence.

Chapitre 1 : Sélection des variables discriminantes

Afin de retenir les variables les plus discriminantes, nous procédons à une étape de segmentation pour regrouper les assurés ayant le même niveau de risque. L'objectif est de sélectionner le minimum de variables explicatives du maintien en incapacité en partant de l'ensemble des variables à notre disposition. Pour ce faire, nous utilisons la méthode des forêts aléatoires.

Le principe de cette méthode consiste à calculer la moyenne des prévisions de plusieurs arbres de régression de type CART²² indépendants pour réduire la variance et donc l'erreur de prévision. Pour chaque arbre, plusieurs échantillons d'individus sont sélectionnés à l'aide de la technique du bootstrap (tirages avec remise). La construction d'un nœud de l'arbre se fait sur un sous-ensemble de variables tirées aléatoirement.

Par rapport à un arbre de régression simple, cette méthode permet d'éviter le sur-apprentissage et d'être plus robuste.

Nous rappelons brièvement l'algorithme de construction d'un arbre de décision (régression ou classification) :

1. Chercher la variable qui sépare le mieux les individus en 2 sous-groupes.
2. Appliquer la séparation à la population.
3. Répéter le même processus pour chaque sous-groupe.
4. S'arrêter lorsque les sous-groupes atteignent une taille minimale ou bien quand le niveau maximal d'homogénéité est atteint.

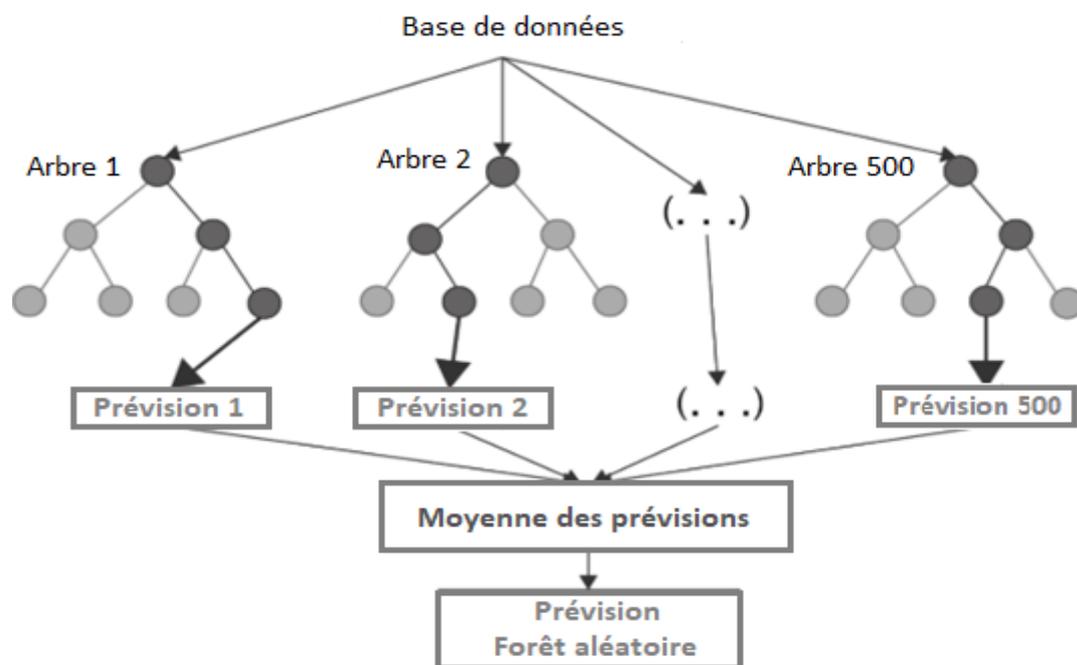


Figure 18- Schématisation de la méthode des forêts aléatoires

²² Se référer au support de cours Introduction aux arbres de décisions (de type CART) de Christophe Chesneau [disponible sur internet : <https://cel.archives-ouvertes.fr/cel-02281064v2>].

L'application de la méthode des forêts aléatoires est réalisée sous le logiciel R en utilisant la fonction *randomForest*. Uniquement les arrêts clos sont pris en compte dans la construction du modèle. La variable à expliquer est la durée de l'arrêt en jours. Les variables explicatives sont l'âge, la franchise, le type d'arrêt, le sexe, la situation familiale, la CSP, le salaire de référence et le secteur d'activité.

Pour évaluer l'importance des variables dans le modèle, la fonction *VarImpPlot* est utilisée :

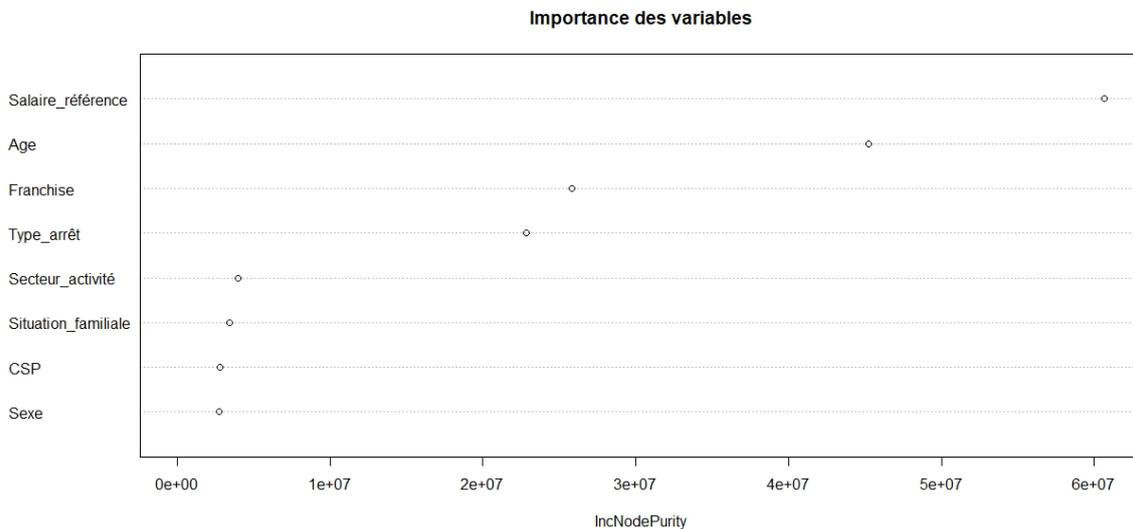


Figure 19- Classement des variables discriminantes avec la méthode des forêts aléatoires

Ce graphique met en évidence que les variables les plus discriminantes sont le salaire de référence, l'âge, la durée de la franchise et le type d'arrêt. Les variables restantes ont moins d'importance dans l'explication de la durée du maintien en incapacité.

Le salaire de référence constitue le facteur le plus important dans l'explication du maintien en incapacité. Le niveau de salaire permet d'apporter une information sur la facilité d'accès aux soins et par conséquent le temps nécessaire pour le rétablissement des pathologies. Ainsi, nous avons pu mettre en exergue une tendance décroissante de la durée moyenne de l'arrêt en fonction du salaire. Toutefois, l'information apportée par la variable salaire de référence est à prendre avec précaution car le niveau de garantie n'est pas forcément identique pour tous les contrats. La segmentation par salaire de référence ne sera donc pas retenue.

Il en ressort aussi que l'âge à la survenance de l'arrêt reste un facteur déterminant quant à la capacité de l'assuré à sortir de l'état d'incapacité. Nous avons vu dans la partie 2 que plus l'âge de l'assuré est élevé, plus la probabilité de reprendre son activité est faible.

La durée de l'arrêt semble être dépendante aussi de la durée de la franchise et du type d'arrêt. Ce constat confirme nos analyses descriptives. Néanmoins, ces variables sont beaucoup moins significatives que l'âge et le salaire de référence. Compte tenu de la composition de nos données²³ et du périmètre de l'étude²⁴, nous faisons le choix de ne pas construire de tables segmentées par franchise et par type d'arrêt.

²³ 90% des arrêts sont causés par une maladie et 77% des observations ont une franchise de 3 jours.

²⁴ La durée de la franchise est limitée à 10 jours sur le portefeuille étudié.

Nous retiendrons pour la suite l'âge à la survenance de l'arrêt comme unique variable explicative de la durée de maintien en incapacité. L'objectif est donc d'obtenir une loi d'expérience à double dimension qui tient compte de l'âge à l'entrée en incapacité et de l'ancienneté dans cet état.

Pour une modélisation robuste, nous devons disposer pour chaque âge à l'arrêt d'un nombre d'observations suffisamment élevé. Nous rappelons dans le graphique ci-dessous la répartition des arrêts par âge :

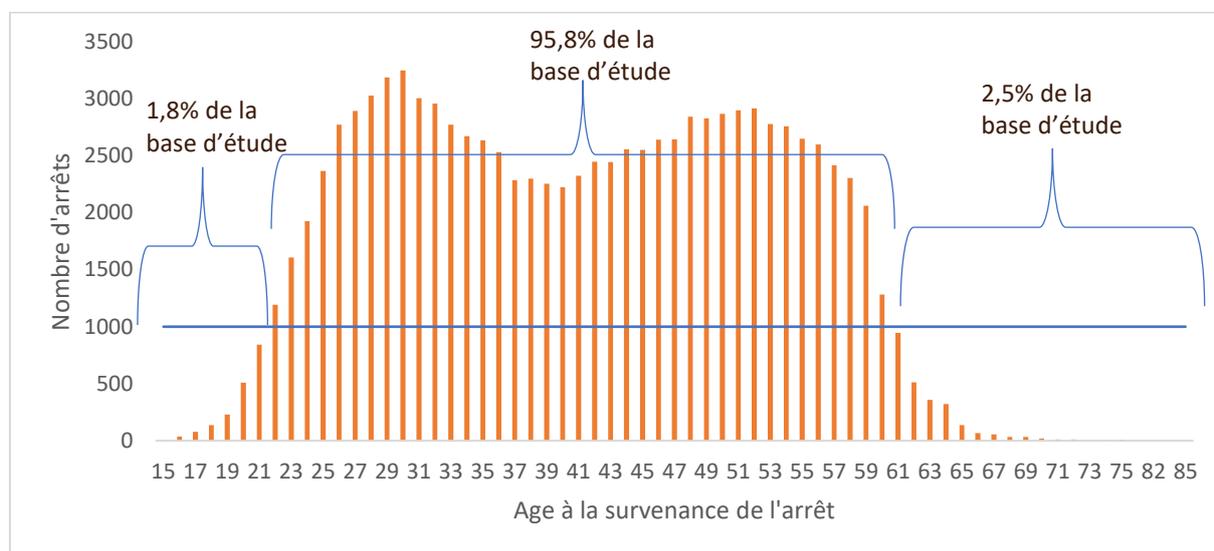


Figure 20- Répartition des arrêts par âge

Environ 96% des arrêts sont concentrés entre 22 ans et 60 ans. Sur cette tranche, nous disposons d'un volume minimum de 1000 arrêts pour chaque âge distinct²⁵.

Etant donné la faible exposition au risque de maintien en incapacité pour les âges supérieurs ou égaux à 61 ans, nous faisons le choix de construire une loi d'expérience limitée à 60 ans. Quant aux âges inférieurs ou égaux à 21 ans, nous choisissons de les regrouper au sein d'une même classe pour avoir un nombre suffisant de personnes soumises au risque de sortie.

Pour rappel, la table de maintien du BCAC est bornée par les âges « 23 ans ou moins » et 66 ans. Une extrapolation de la table pour les âges compris entre 61 ans et 66 ans sera alors proposée ultérieurement (chapitre 6).

Chapitre 2 : Choix de la période d'observation

Nous rappelons que la période d'observation des données s'étale sur 16 ans : du 01/01/2004 au 31/12/2019. Cette profondeur historique permet d'avoir une bonne volumétrie des données, tout en vérifiant la stabilité de la durée de maintien en incapacité sur cette période (cf. Partie 2 : Analyse descriptive des données). La base d'étude retenue pour la modélisation comporte 103 010 observations relatives à des arrêts survenus entre le 01/01/2004 et le 31/12/2018.

²⁵ Seuil minimum retenu pour avoir une modélisation « fine » par âge.

L'observation des données après un recul d'un an par rapport à la dernière date de survenance permet d'éviter les problématiques de déclarations tardives et d'avoir des données majoritairement complètes. Par ailleurs, afin de tester la robustesse de notre modélisation, les calculs et les comparaisons des PM seront effectués sur les sinistres de la survenance 2019²⁶.

Chapitre 3 : Présentation des notions de censure et de troncature

La construction d'une loi de maintien en incapacité se base sur l'observation de la durée des arrêts pour l'ensemble des sinistrés. Cependant, pour certains individus, tout ou une partie de l'information n'est pas observable : on parle alors de données censurées ou tronquées. Ces notions sont explicitées ci-après :

Censure :

Dans notre étude, la censure correspond à la présence d'un arrêt dans la base de données, mais avec une information incomplète au regard de sa durée.

On parle de censure à droite lorsque la sortie de l'état d'incapacité se produit après la fin de la période d'observation. Dans notre cas, tous les arrêts qui sont en cours au 31/12/2019 sont concernés par une censure à droite. Ces données représentent moins de 1% de la base totale.

On parle de censure à gauche lorsque l'entrée en état d'incapacité se produit avant le début de la période d'observation. Dans notre cas, nous disposons de tout l'historique des assurés présents dans le portefeuille et qui ont subi un sinistre incapacité. Il n'y a pas donc de censure à gauche.

Troncature :

Dans notre étude, la troncature correspond à l'absence d'un arrêt dans la base de données alors que celui-ci s'est en réalité produit.

On parle de troncature à gauche lorsque l'arrêt n'est pas répertorié dans la base de données en dessous d'une certaine durée. La franchise peut être considérée comme une troncature à gauche. En effet, l'assureur n'a connaissance que des sinistres qui ont dépassé la durée de la franchise.

On parle de troncature à droite lorsque l'arrêt n'est pas répertorié dans la base de données s'il dépasse une certaine durée. Ce type de troncature n'est pas présent dans notre cas.

Remarques :

- La censure à droite et la troncature à gauche sont plus courantes dans les situations d'assurance.
- Contrairement aux données censurées qui sont partiellement observables, l'information sur les données tronquées est entièrement perdue.
- Les censures et les troncatures affectent l'estimation de la loi de maintien en incapacité. L'estimateur utilisé pour la modélisation doit prendre en compte cette particularité.

²⁶ Données qui n'ont pas servis à la modélisation.

Chapitre 4 : Estimation des taux bruts de sortie d'incapacité

Ne connaissant pas à priori l'allure de la loi à modéliser, notre choix s'oriente vers une méthode non paramétrique pour estimer les taux bruts de sortie d'incapacité. L'une des méthodes les plus couramment utilisées dans la pratique est celle de Kaplan-Meier. L'avantage de cette méthode est sa capacité à tenir compte de données incomplètes et plus particulièrement des censures à droite.

1. Présentation théorique de l'estimateur de Kaplan-Meier

L'estimateur de Kaplan-Meier est une méthode non paramétrique d'estimation de la fonction de survie. Il a été introduit en 1958 par Edward Kaplan et Paul Meier dans un article publié dans le journal de l'*American Statistical Association*.

Le principe de cet estimateur est le suivant : la période de survie est discrétisée en sous-périodes délimitées par des dates. La probabilité d'un assuré de ne pas sortir de l'état d'incapacité lors d'une sous-période est équivalente au fait de ne pas sortir d'incapacité juste avant cette sous-période et de ne pas sortir d'incapacité lors de cette sous-période. Ainsi, la probabilité de ne pas sortir d'incapacité à une sous-période donnée peut se calculer conditionnellement au fait de ne pas sortir d'incapacité juste avant cette sous-période.

Le temps entre deux instants t et $t+1$ est ainsi discrétisé en n sous-intervalles de temps, dont la forme est $[t_i; t_{i+1}[, 1 \leq i \leq n$.

Notations :

x : Age d'entrée en incapacité

T : Durée de maintien en incapacité au-delà de la date t

$q_{x,i}$: Probabilité d'un assuré d'âge x de sortir de l'état d'incapacité à la date t_i , soit

$$q_{x,i} = \mathbb{P}(T < t_{i+1} | T \geq t_i)$$

$S_x(t)$: Fonction de survie au point t pour un assuré d'âge x , soit $S_x(t) = \mathbb{P}(T > t)$

$n_{x,i}$: Population d'âge x sous risque juste avant la date t_i

$d_{x,i}$: Nombre de sorties d'incapacité en t_i parmi une population d'âge x

$c_{x,i}$: Nombre de censures à droite en t_i parmi une population d'âge x

A noter que $n_{x,i}$ s'obtient en utilisant la formule de récurrence suivante :

$$n_{x,i} = n_{x,i-1} - d_{x,i-1} - c_{x,i-1}$$

La fonction de survie définie précédemment peut s'exprimer de la façon suivante :

$$S_x(t) = \prod_{i|t_i < t} (1 - q_{x,i}), \forall t \in [0, t_n]$$

La méthode du maximum de vraisemblance est utilisée pour estimer les $q_{x,i}$ à partir de l'observation d'un n-échantillon.

Le nombre de sorties d'incapacité suit une loi binomiale $\mathcal{B}(n_{x,i}; q_{x,i})$. La fonction de vraisemblance s'écrit alors comme suit :

$$\mathcal{L}(q_{x,1}, q_{x,2}, \dots, q_{x,n}) = \prod_{i=1}^n C_{n_{x,i}}^{d_{x,i}} q_{x,i}^{d_{x,i}} (1 - q_{x,i})^{n_{x,i} - d_{x,i}}$$

Trouver le maximum de \mathcal{L} revient à trouver la solution de l'équation :

$$\frac{\partial \ln(\mathcal{L})}{\partial q_{x,i}} = 0, \forall i \in [1, n]$$

$$\Leftrightarrow \frac{\partial \sum_{i=1}^n (\ln(C_{n_{x,i}}^{d_{x,i}}) + d_{x,i} \ln(q_{x,i}) + (n_{x,i} - d_{x,i}) \ln(1 - q_{x,i}))}{\partial q_{x,i}} = 0, \forall i \in [1, n]$$

$$\Leftrightarrow \frac{d_{x,i}}{\hat{q}_{x,i}} - \frac{n_{x,i} - d_{x,i}}{1 - \hat{q}_{x,i}} = 0, \forall i \in [1, n]$$

$$\Leftrightarrow \hat{q}_{x,i} = \frac{d_{x,i}}{n_{x,i}}, \forall i \in [1, n]$$

Il s'en déduit alors l'estimateur de la fonction de survie :

$$\hat{S}_x(t) = \prod_{i|t_i < t} (1 - \frac{d_{x,i}}{n_{x,i}}), \forall t \in [0, t_n]$$

Cette formule repose sur l'hypothèse d'absence d'ex aequo qui signifie que la probabilité que deux personnes sortent exactement au même instant de l'état d'incapacité est nulle presque sûrement. La validité de cette hypothèse est remise en question dans notre cas en raison de la discrétisation journalière du temps.

Pour y remédier, la formule simplifiée suivante est utilisée :

$$\hat{S}_x(t) = \prod_{i=AncMin}^t (1 - \frac{d_{x,i}}{n_{x,i}}), \forall t \in \{AncMin, \dots, AncMax\} \quad (1)$$

Où *AncMin* et *AncMax* représentent respectivement les anciennetés minimale et maximale en état d'incapacité sur la population étudiée.

Une fois l'estimateur de la fonction de survie calculé, il est possible d'en déduire les taux de sortie bruts $\hat{q}_{x,t}$ pour un âge x et une ancienneté t fixés :

$$\hat{q}_{x,t} = 1 - \frac{\hat{S}_x(t+1)}{\hat{S}_x(t)}$$

Remarques :

- 1- Si aucune censure n'intervient, l'estimateur de Kaplan-Meier correspond à la fonction de répartition empirique.

2- L'estimateur de Kaplan-Meier présente les propriétés suivantes²⁷ :

- Il est cohérent : $\hat{S}_x(t) = \frac{1}{n} \sum_{i=1}^n 1_{\{t_i > t\}} + \sum_{i=1}^n 1_{\{t_i \leq t, d_i = 0\}} \frac{\hat{S}_x(t)}{\hat{S}_x(t_i)}$;
- Il est convergent : $\forall \varepsilon > 0, \lim_{n \rightarrow \infty} \mathbb{P}(|\hat{S}_x(t) - S_x(t)| \geq \varepsilon) = 0$;
- Il est asymptotiquement gaussien : $\sqrt{n_{x,i}}(\hat{S}_x(t) - S_x(t)) \xrightarrow{loi} W$, avec W un processus gaussien centré ;
- Il est biaisé positivement : $\mathbb{E}(\hat{S}_x(t)) \geq S_x(t)$.

3- Plusieurs estimateurs existent pour approcher la variance de l'estimateur de Kaplan-Meier. L'un des plus connus est l'estimateur de Greenwood :

$$\mathbb{V}(\hat{S}_x(t)) = (\hat{S}_x(t))^2 \sum_{i=AncMin}^t \frac{d_{x,i}}{n_{x,i} * (n_{x,i} - d_{x,i})}$$

Cette expression est obtenue en utilisant l'approximation suivante (méthode delta) : pour une fonction f régulière et une variable aléatoire $X \approx \mu + \sigma Z$ (où σ est petit et Z une variable aléatoire centrée réduite), nous avons le développement limité suivant :

$$\mathbb{V}(f(X)) = \mathbb{V}(f(\mu + \sigma Z)) \approx \mathbb{V}(f(\mu) + \sigma Z f'(\mu)) = \sigma^2 (f'(\mu))^2$$

Ainsi, en prenant $f = \ln$ et $X = \hat{S}_x(t)$:

$$\mathbb{V}(\ln(\hat{S}_x(t))) \approx \mathbb{V}(\hat{S}_x(t)) * \left(\frac{1}{\hat{S}_x(t)}\right)^2 \quad (2)$$

D'autre part, d'après (1) nous pouvons écrire :

$$\ln(\hat{S}_x(t)) = \sum_{i=AncMin}^t \ln(1 - \hat{q}_{x,i}), \forall t \in \{AncMin, \dots, AncMax\}$$

Comme $n_{x,i}(1 - \hat{q}_{x,i})$ suit une loi binomiale de paramètres $(n_{x,i}, (1 - \hat{q}_{x,i}))$, alors en appliquant la méthode delta :

$$\begin{aligned} \mathbb{V}(\ln(1 - \hat{q}_{x,i})) &= \mathbb{V}(\ln(n_{x,i}(1 - \hat{q}_{x,i}))) \\ &\approx \left(\frac{1}{n_{x,i}(1 - \hat{q}_{x,i})}\right)^2 n_{x,i}(1 - \hat{q}_{x,i})\hat{q}_{x,i} \\ &= \frac{\hat{q}_{x,i}}{n_{x,i}(1 - \hat{q}_{x,i})} \\ &= \frac{d_{x,i}}{n_{x,i}(n_{x,i} - d_{x,i})} \end{aligned}$$

En supposant l'indépendance des $\hat{q}_{x,i}$, nous pouvons écrire :

²⁷ Pour plus de détails sur les propriétés de l'estimateur de Kaplan-Meier, se référer au support de cours de modèles de durée de Frédéric Planchet sur les statistiques des modèles non paramétriques [Disponible sur Internet : <http://www.ressources-actuarielles.net>].

$$\mathbb{V}\left(\ln\left(\hat{S}_x(t)\right)\right) = \mathbb{V}\left(\sum_{i=AnclMin}^t \ln(1 - \hat{q}_{x,i})\right) = \sum_{i=AnclMin}^t \frac{d_{x,i}}{n_{x,i}(n_{x,i} - d_{x,i})} \quad (3)$$

$$(2) + (3) \Rightarrow \mathbb{V}\left(\hat{S}_x(t)\right) = \left(\hat{S}_x(t)\right)^2 \sum_{i=AnclMin}^t \frac{d_{x,i}}{n_{x,i}(n_{x,i} - d_{x,i})}$$

- 4- L'estimateur de Kaplan-Meier présenté ci-avant correspond à un cadre monodimensionnel. C'est-à-dire un taux de sortie par âge à un instant donné. Des travaux relatifs à une estimation bidimensionnelle ont été menés afin d'obtenir un taux par âge à plusieurs dates. Cependant, l'estimation des taux bruts âge par âge par Kaplan-Meier reste la méthode la plus robuste. En effet, « la perte d'information que représente la non prise en compte de la loi conjointe selon les deux dimensions du problème est faible et peu pénalisante en pratique »²⁸.

2. Analyse des taux de sortie d'incapacité obtenus

Nous représentons ci-après les résultats issus de l'application de la méthode de Kaplan-Meier. La modélisation ponctuelle des taux bruts de sortie a été réalisée avec le logiciel R en utilisant la fonction *survfit* du package *survival*. Celle-ci indique également les intervalles de confiance à 95% en se basant sur l'estimateur de Greenwood.

Bien que les calculs soient réalisés avec un pas journalier, seuls les résultats mensuels seront retenus pour la construction de la loi de maintien en incapacité. Nous retrouverons ainsi la même structure de loi que celle utilisée dans la table règlementaire du BCAC.

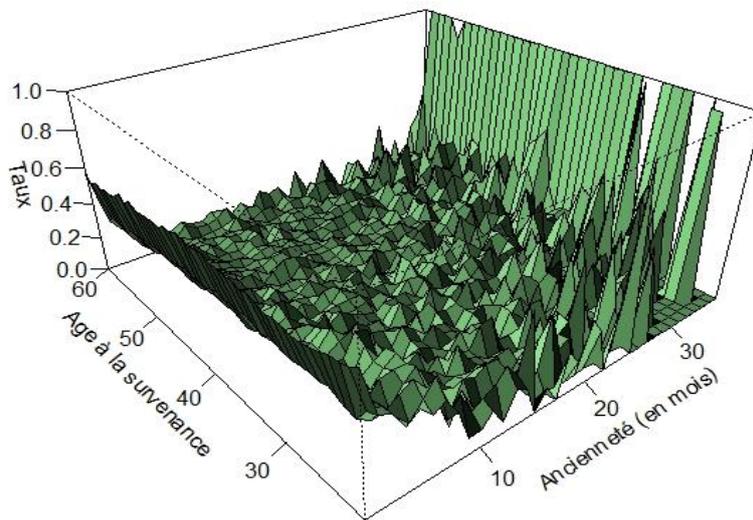


Figure 21- Taux de sortie bruts sur 36 mois

Une grande volatilité est observée pour les anciennetés les plus élevées surtout lorsque l'âge d'entrée en incapacité est faible. Ce résultat était attendu au vu de la volumétrie disponible (Pour rappel 97% des sinistres sont clos dès la 1^{ère} année). Pour mieux apprécier la précision des résultats, une

²⁸ PLANCHET F., THEROND P., Modèles de Durée - Applications actuarielles, Economica, 2006.

représentation de la taille de l'intervalle de confiance à 95% des taux de sortie est donnée dans le graphique ci-dessous :

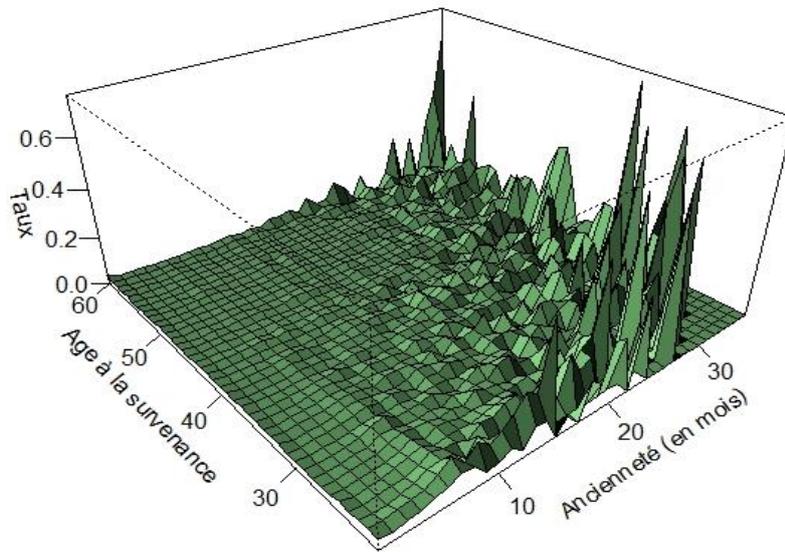


Figure 22- Taille de l'intervalle de confiance à 95% des taux de sortie

Les résultats semblent être globalement peu volatiles pour les anciennetés inférieures à 12 mois. La taille de l'intervalle de confiance augmente considérablement au-delà et notamment pour les âges les moins élevés car nous avons peu de données sur ce segment.

Pour éviter de biaiser les résultats à cause de la faible volumétrie, nous considérerons par la suite les anciennetés en incapacité jusqu'à 12 mois. Les taux de sortie pour les anciennetés au-delà de ce seuil seront ensuite déterminés par extrapolation.

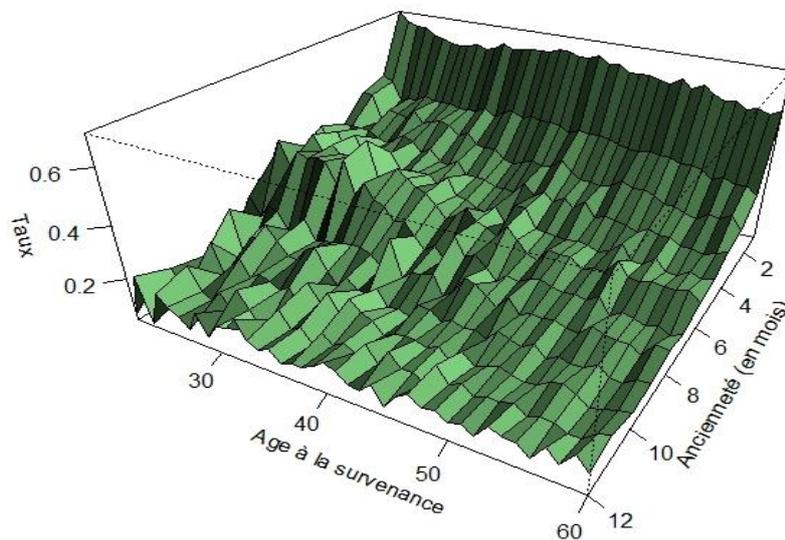


Figure 23- Taux de sortie bruts sur 12 mois

En limitant l'ancienneté à 12 mois, nous obtenons une surface plus régulière. De plus, une tendance décroissante des taux de sortie avec l'ancienneté en arrêt ressort : plus l'ancienneté en arrêt augmente, plus la probabilité de sortir de l'état d'incapacité diminue et donc la probabilité d'en rester croît. Ce résultat, également observé avec la loi de maintien en incapacité du BCAC, peut s'expliquer par la dégradation de l'état de santé due à l'évolution pathologique au cours du temps.

Pour mieux visualiser les résultats obtenus, nous exposons les taux de sortie bruts et les intervalles de confiance associés pour différents âges d'entrée en incapacité :

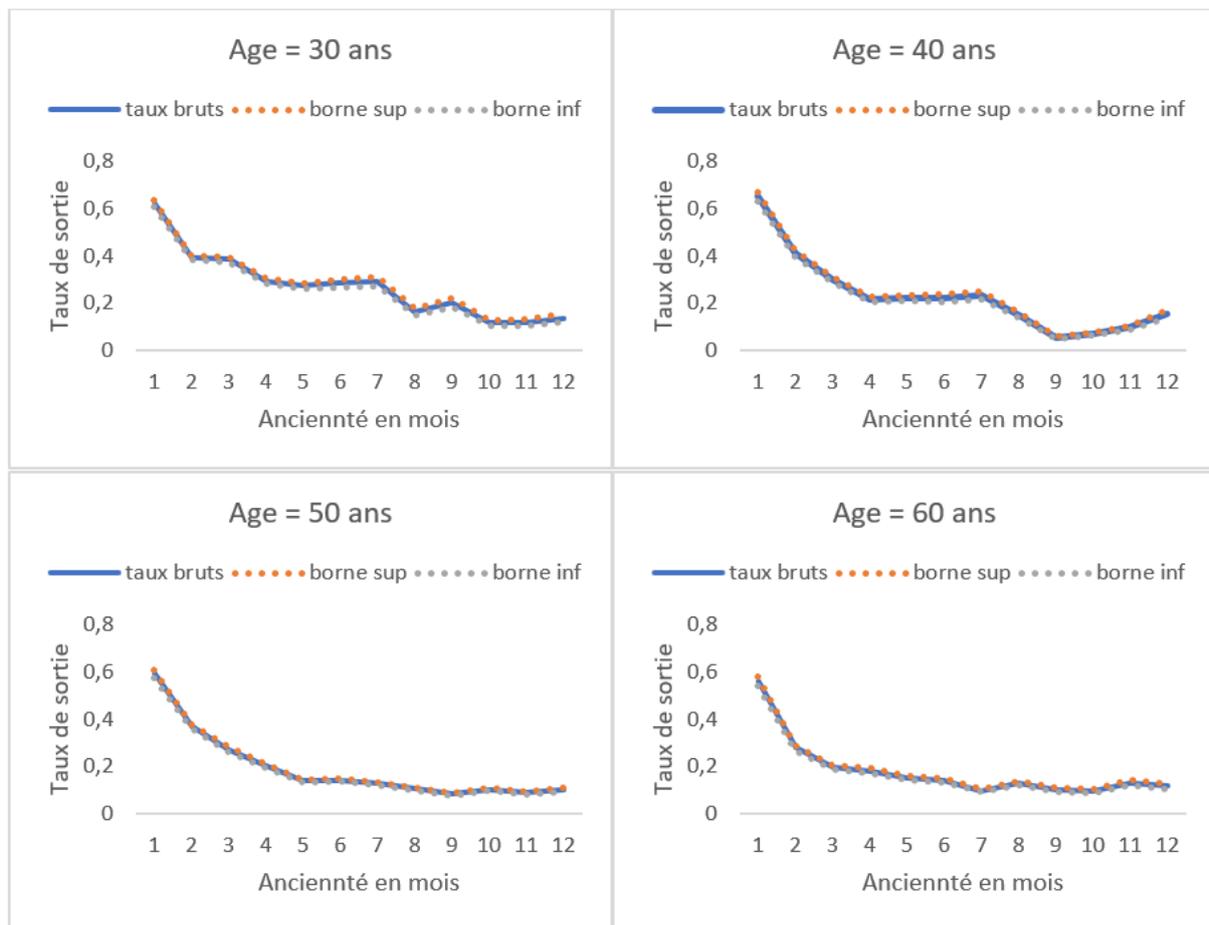


Figure 24- Taux de sortie bruts sur 12 mois : zoom par âge

La tendance décroissante des taux de sortie avec l'ancienneté est mise en évidence graphiquement. La fiabilité du modèle est validée au vu de la taille de l'intervalle de confiance pour chaque âge représenté : l'écart le plus important est situé au niveau de l'âge d'entrée en incapacité de 30 ans pour une ancienneté de 7 mois (taux brut = 0,29, borne sup = 0,31 et borne inf = 0,27).

Par ailleurs, certaines variations assez irrégulières sont remarquées notamment pour les âges les moins élevés. Un lissage des taux de sortie bruts est donc nécessaire pour les diminuer.

Chapitre 5 : Lissage des taux de sortie d'incapacité

Dans cette partie nous allons nous intéresser au lissage des taux de sortie bruts obtenus. Le but de cette étape est de rendre les taux modélisés plus réguliers tout en restant proche de la réalité. Pour ce faire, deux contraintes doivent être prises en compte de manière conjointe :

- La fidélité : les taux ajustés doivent être relativement proches des taux bruts ;
- La régularité : la courbe des taux ajustés doit être la plus lisse possible.

Comme pour l'estimation des taux bruts, nous optons pour une méthode de lissage non paramétrique car nous ne connaissons pas a priori la loi théorique à laquelle s'adapter. L'une des méthodes les plus utilisées dans la pratique est celle de Whittaker-Henderson.

1. Présentation théorique de la méthode de Whittaker-Henderson

Le principe de la méthode de Whittaker-Henderson est de combiner linéairement la fidélité et la régularité de façon à minimiser la somme de ces deux critères. Le but est donc d'obtenir le meilleur compromis fidélité/régularité.

Notations :

F : Critère de fidélité

R : Critère de régularité

$q = (q_i)_{1 \leq i \leq n}$: Vecteur des taux lissés

$\hat{q} = (\hat{q}_i)_{1 \leq i \leq n}$: Vecteur des taux bruts

$w = (w_i)_{1 \leq i \leq n}$: Vecteur de pondération

$W = \text{diag}(w_i)_{1 \leq i \leq n}$: Matrice diagonale des poids $(w_i)_{1 \leq i \leq n}$

Le critère de fidélité est défini de la manière suivante :

$$F = \sum_{i=1}^n w_i (q_i - \hat{q}_i)^2$$

Avec n le nombre de modalités considérées.

Ce critère mesure la distance euclidienne entre les taux bruts et les taux lissés. Chaque distance est pondérée par le poids w_i . En pratique, les poids sont égaux soit à 1, soit au rapport entre l'effectif associé au taux considéré et l'effectif total. La seconde pondération permet de limiter l'importance des taux bruts pour les anciennetés avec un nombre faible d'observations.

Quant au critère de régularité, il est défini de la façon suivante :

$$R = \sum_{i=1}^{n-z} (\Delta^z q_i)^2$$

Avec Δ^z l'opérateur de différenciation d'ordre z :

$$\Delta^z q_i = \sum_{k=0}^z \binom{z}{k} (-1)^{z-k} q_{i+k}, \forall i \in \{1, \dots, n\}$$

$$\text{Par définition : } \left\{ \begin{array}{l} \Delta q_i = q_{i+1} - q_i \\ \Delta^2 q_i = \Delta(\Delta q_i) \\ \quad = \Delta(q_{i+1} - q_i) \\ \quad = (q_{i+2} - q_{i+1}) - (q_{i+1} - q_i) \\ \quad = q_{i+2} - 2q_{i+1} + q_i \\ \quad \dots \\ \Delta^z q_i = \Delta(\Delta^{z-1} q_i) \\ \quad = \Delta(\Delta(\Delta^{z-2} q_i)) \\ \quad = \dots \\ \quad = \sum_{k=0}^z \binom{z}{k} (-1)^{z-k} q_{i+k} \end{array} \right.$$

Le paramètre z permet de contrôler la régularité du lissage.

Si $z = 1$, nous retrouvons la distance euclidienne classique entre deux valeurs lissées successives.

Pour tenir compte de suffisamment d'informations consécutives, le paramètre z est souvent compris entre 2 et 5.

Une fois les deux critères définis, il faut alors minimiser une combinaison linéaire de F et de R , qui prend la forme :

$$M = F + hR$$

Où h est un paramètre qui permet de contrôler le poids associé à chacun des deux termes.

Remarquons que si $h = 0$, $M = F$, le lissage est alors nul. En d'autres termes, les taux qui sont les plus fidèles aux taux bruts sont les taux bruts eux-mêmes.

Pour minimiser M , il faut alors résoudre le système d'équations suivant :

$$\frac{\partial M}{\partial q_i} = 0, \forall i \in \{1, \dots, n\}$$

La résolution de ce système peut se faire en utilisant les notations matricielles exposées ci-dessus.

$$\text{Il vient alors : } F = (q - \hat{q})^T W (q - \hat{q})$$

$$\text{Et } R = (\Delta^z q)^T (\Delta^z q)$$

$$\text{Où } \Delta^z q = (\Delta^z q_i)_{1 \leq i \leq n-z}$$

Par ailleurs, en notant K_z la matrice de dimension $(n - z) \times n$, avec pour termes les coefficients binomiaux d'ordre z dont le signe alterne et commence positivement lorsque z est pair, il s'avère que :

$$\Delta^z q = K_z q$$

Par exemple, pour $n = 5$ et $z = 2$,

$$K_2 = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \end{pmatrix}$$

Et pour $n = 5$ et $z = 3$,

$$K_3 = \begin{pmatrix} -1 & 3 & -3 & 1 & 0 \\ 0 & -1 & 3 & -3 & 1 \end{pmatrix}$$

Il résulte que :

$$\begin{aligned} M &= (q - \hat{q})^T W (q - \hat{q}) + h q^T K_z^T K_z q \\ &= q^T W q - 2 q^T W \hat{q} + \hat{q}^T W \hat{q} + h q^T K_z^T K_z q \end{aligned}$$

Car $q^T W \hat{q} = \hat{q}^T W q$

Par ailleurs, $\frac{\partial (q^T W q)}{\partial q} = (W + W^T) q = 2W q$ (car W est une matrice diagonale donc symétrique)

Et $\frac{\partial (q^T K_z^T K_z q)}{\partial q} = 2K_z^T K_z q$ (car $K_z^T K_z$ est symétrique par propriété)

Ce qui donne :

$$\begin{aligned} \frac{\partial M}{\partial q} &= 2W q - 2W \hat{q} + 2h K_z^T K_z q \\ &= 2(W + h K_z^T K_z) q - 2W \hat{q} \end{aligned}$$

En supposant que la matrice $W + h K_z^T K_z$ est inversible, cette dérivée s'annule pour :

$$q = (W + h K_z^T K_z)^{-1} W \hat{q}$$

Remarque : En pratique, la décomposition de Cholesky²⁹ peut être utilisée pour réaliser l'inversion de la matrice symétrique positive $W + h K_z^T K_z$.

Extension en dimension 2 :

Le lissage décrit ci-dessus présente un intérêt lorsque les données traitées sont unidimensionnelles (exemple : table de mortalité par âge). Cependant, la problématique du maintien en incapacité est bidimensionnelle. D'où l'intérêt de la généralisation en dimension 2 de la méthode de Whittaker-Henderson. Cela permet un lissage conjoint dans les deux directions, plus efficace que le lissage séparé selon chaque variable. Le principe de cette généralisation est de se ramener au cas unidimensionnel.

Nous disposons des estimations des taux bruts $\hat{q} = (\hat{q}_i)_{1 \leq i \leq m, 1 \leq j \leq n}$. Le critère de fidélité se généralise simplement de la manière suivante :

$$F = \sum_{i=1}^m \sum_{j=1}^n w_{ij} (q_{ij} - \hat{q}_{ij})^2$$

L'extension en dimension 2 du critère de régularité est un peu plus délicate. Pour cela nous distinguons d'abord la régularité verticale via l'opérateur Δ_v^z (qui agit sur les q_{ij} à j fixé vu comme une série indexée par i) et la régularité horizontale via l'opérateur Δ_h^z (qui agit sur les q_{ij} à i fixé vu comme une série indexée par j).

Ainsi un indice de régularité verticale est calculé comme suit :

$$R_v = \sum_{j=1}^n \sum_{i=1}^{m-z} (\Delta_v^z q_{ij})^2$$

²⁹ Théorème : Si A est une matrice symétrique définie positive, il existe une matrice réelle triangulaire inférieure L telle que $A = LL^T$.

De la même manière un indice de régularité horizontale est défini comme suit :

$$R_h = \sum_{i=1}^m \sum_{j=1}^{n-z} (\Delta_h^y q_{ij})^2$$

Par la suite, nous devons minimiser :

$$M = F + \alpha R_v + \beta R_h$$

La résolution de ce problème d'optimisation s'effectue en réarrangeant les différents éléments pour se ramener au cas unidimensionnel. Pour cela, un vecteur u de taille $m \times n$ est défini tel que $u_{n(i-1)+j} = \hat{q}_{ij}$ (cela revient à prendre pour les n premiers éléments du vecteur u la première ligne de la matrice \hat{q} , puis les éléments de la seconde ligne et ainsi de suite). De même, une matrice de poids est construite en copiant sur la diagonale les lignes de la matrice W ($W_{n(i-1)+j, n(i-1)+j} = W_{ij}$). Enfin, les matrices K_v^z et K_h^y sont définies en procédant de la même manière.

Les taux lissés s'obtiennent alors par :

$$q^* = (W^* + \alpha K_v^z{}^T K_v^z + \beta K_h^y{}^T K_h^y)^{-1} W^* u$$

2. Description des tests d'adéquation de l'ajustement

Une fois le lissage effectué, nous devons vérifier la validité de l'ajustement afin de s'assurer notamment que les taux lissés ne soient pas trop éloignés des taux bruts. Pour cela, nous appliquons le test du χ^2 . Ce test permet de mesurer la distance entre les taux révisés et les taux d'origine.

La statistique du χ^2 est définie par :

$$Z = \sum_{k=AncMin}^{AncMax} \sum_{x=AgeMin}^{AgeMax} n_{x,k} \frac{(q_{x,k}^{brut} - q_{x,k}^{lissé})^2}{q_{x,k}^{lissé}}$$

Où

$n_{x,k}$: Exposition pour un âge x et une ancienneté en incapacité de k mois

$q_{x,k}^{lissé}$: Taux de maintien lissé pour un assuré d'âge x et d'ancienneté en incapacité de k mois

$q_{x,k}^{brut}$: Taux de maintien brut pour un assuré d'âge x et d'ancienneté en incapacité de k mois

La statistique Z suit asymptotiquement une loi du χ^2 à $(AncMax - AncMin) + (AgeMax - AgeMin)$ degrés de libertés. Cette statistique est comparée avec le quantile de la loi du χ^2 d'ordre $1 - \alpha$ noté c_α avec α un seuil de confiance usuellement fixé à 5% ou 1%.

Si, $Z \geq c_\alpha$, alors l'adéquation de l'ajustement des taux bruts par le lissage est rejetée.

La qualité de l'ajustement est également évaluée en étudiant les critères de régularité et de fidélité à travers les formules suivantes :

$$Fidélité : F = \sum_{k=AncMin}^{AncMax} \sum_{x=AgeMin}^{AgeMax} (q_{x,k}^{brut} - q_{x,k}^{lissé})^2$$

$$\text{Régularité} : R = \sum_{k=\text{AncMin}}^{\text{AncMax}} \sum_{x=\text{AgeMin}}^{\text{AgeMax}} (q_{x,k}^{\text{lissé}} - q_{x+1,k}^{\text{lissé}})^2$$

Plus ces quantités tendent vers 0, plus le lissage est jugé de bonne qualité.

3. Application de la méthode de Whittaker-Henderson aux données et validation de l'ajustement

Nous affichons ci-après les résultats issus de l'application de la méthode de Whittaker-Henderson à l'aide du logiciel R³⁰.

Nous avons testé plusieurs valeurs des paramètres de lissage et nous avons retenu celles qui assurent un bon compromis entre la fidélité et la régularité.

Nous avons retenu $z = y = 4$. Pour rappel, les paramètres z et y représentent respectivement le degré de régularité verticale et le degré de régularité horizontale.

Pour calibrer l'intensité du lissage, les paramètres qui ont été retenues sont : $\alpha = \beta = 1$. Pour rappel, les paramètres α et β permettent de pondérer respectivement le critère de régularité verticale et le critère de régularité horizontale.

Quant à la matrice des poids, nous avons choisi d'utiliser la pondération par l'exposition des effectifs.

Nous obtenons ainsi la courbe des taux de sortie suivante :

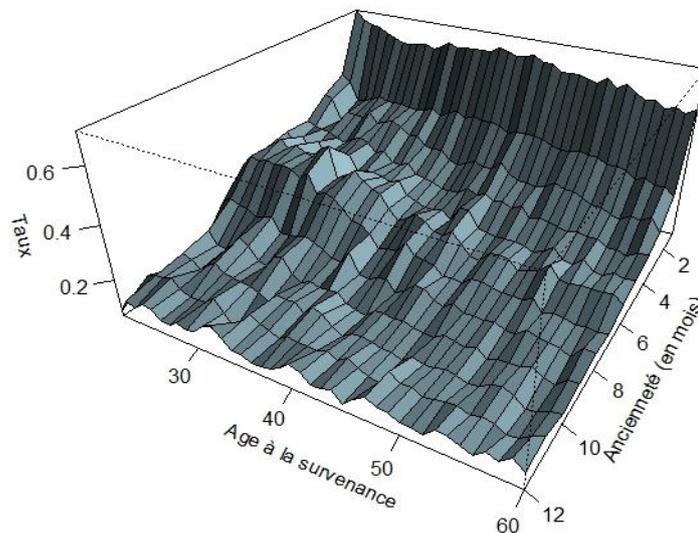


Figure 25- Taux de sortie lissés sur 12 mois

Il est clair que la surface des taux de sortie lissés présente moins d'irrégularités que celle des taux de sortie bruts tout en restant proche de la réalité. Pour valider ce lissage, le test du χ^2 et une étude des critères de fidélité et de régularité sont effectués. Un seuil de confiance $\alpha = 5\%$ est retenu. La loi du χ^2 à 50 degrés de liberté³¹ donne une valeur c_α de 67,50.

³⁰ Nous avons exploité les codes du lissage mis à disposition sur le site *Ressources-Actuarielles*.

³¹ Nombre de degrés de liberté = $(12-1) + (60-21) = 50$.

$$Z = 66,71$$

$$F = 0,13$$

$$R = 0,26$$

L'hypothèse d'adéquation du lissage aux taux bruts est vérifiée puisque $Z < c_\alpha$. De plus, au vu des valeurs des critères de fidélité et de régularité, l'ajustement semble satisfaisant.

Pour apprécier davantage la qualité du lissage, les taux lissés et les taux bruts ont été superposés dans la figure suivante :

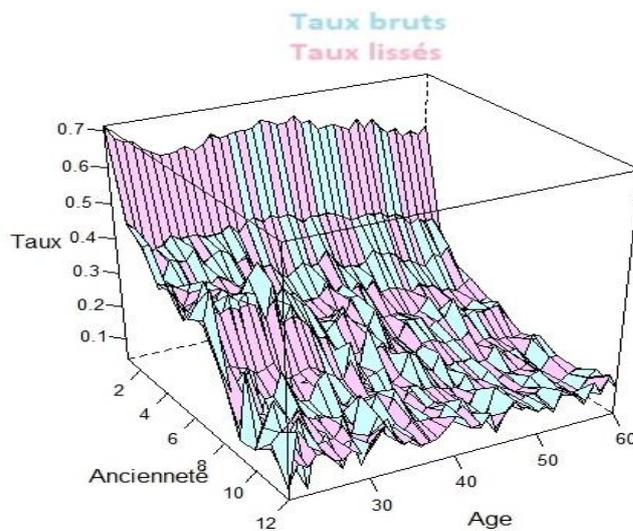


Figure 26- Comparaison des taux de sortie bruts et lissés

Les taux bruts et les taux lissés sont très proches. Pour mieux visualiser les résultats obtenus, nous représentons les taux de sortie lissés et les intervalles de confiance associés aux taux bruts pour différents âges d'entrée en incapacité :

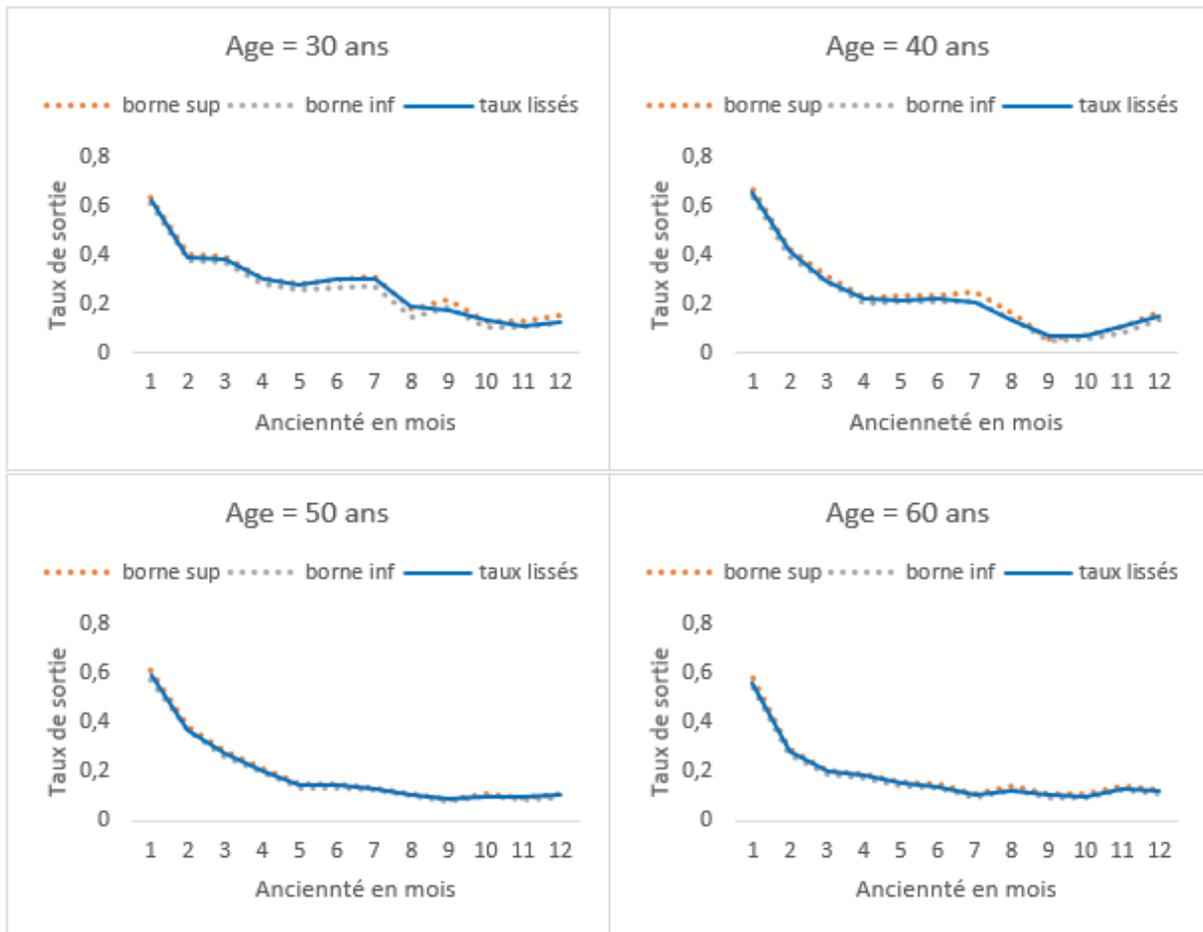


Figure 27- Taux lissés sur 12 mois : zoom par âge

Les taux lissés sont bien compris dans l'intervalle de confiance à 95% des taux bruts issus du modèle de Kaplan-Meier. Ce qui constitue un indicateur de bonne qualité de l'ajustement.

Une fois l'ajustement validé statistiquement et graphiquement, il est intéressant de vérifier l'adéquation du modèle aux données du portefeuille. Pour cela nous comparons le nombre de sorties d'incapacité observées et le nombre de sorties d'incapacité prédites par le modèle.

Le nombre de sorties observées correspond au comptage des sorties réelles mois par mois constatées sur l'ensemble de la base d'étude. Le nombre de sorties prédites correspond à la multiplication de l'effectif sous risque au début de chaque mois par le taux de sortie donné par le modèle.

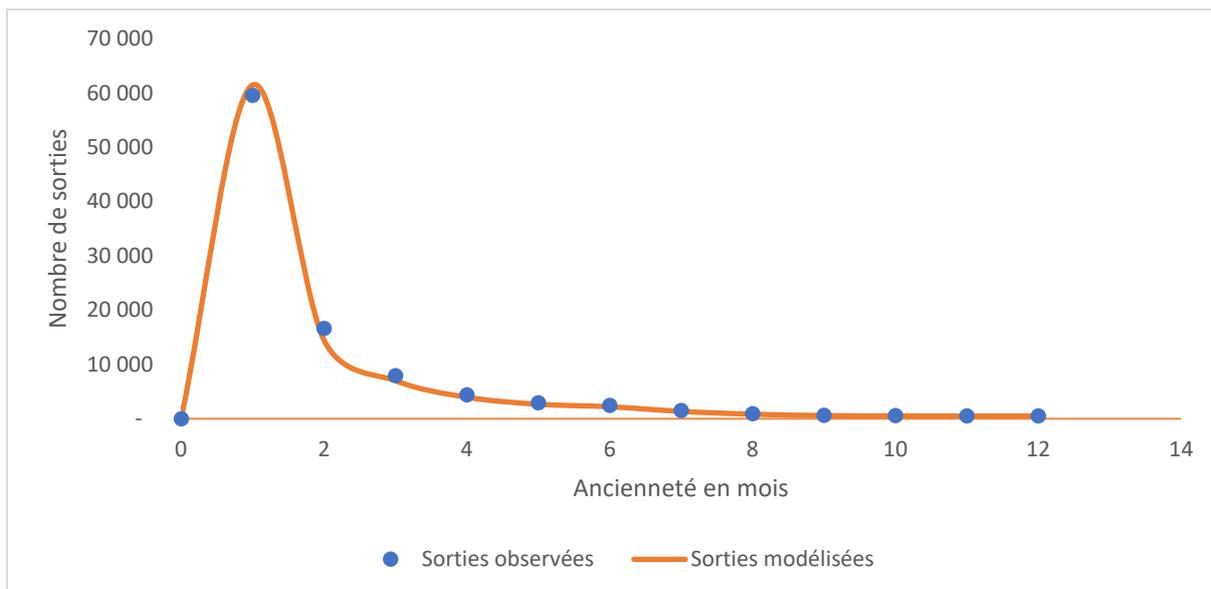


Figure 28- Sorties modélisées vs. Sorties observées

Le modèle semble au global bien prédire les sorties d'incapacité avec un ratio de sorties modélisées par rapport aux sorties observées de 97,2%. Le principe de prudence est ainsi respecté.

Une comparaison graphique est également réalisée pour les sorties relatives à des assurés d'âge d'entrée en incapacité de 30, 40, 50 et 60 ans.

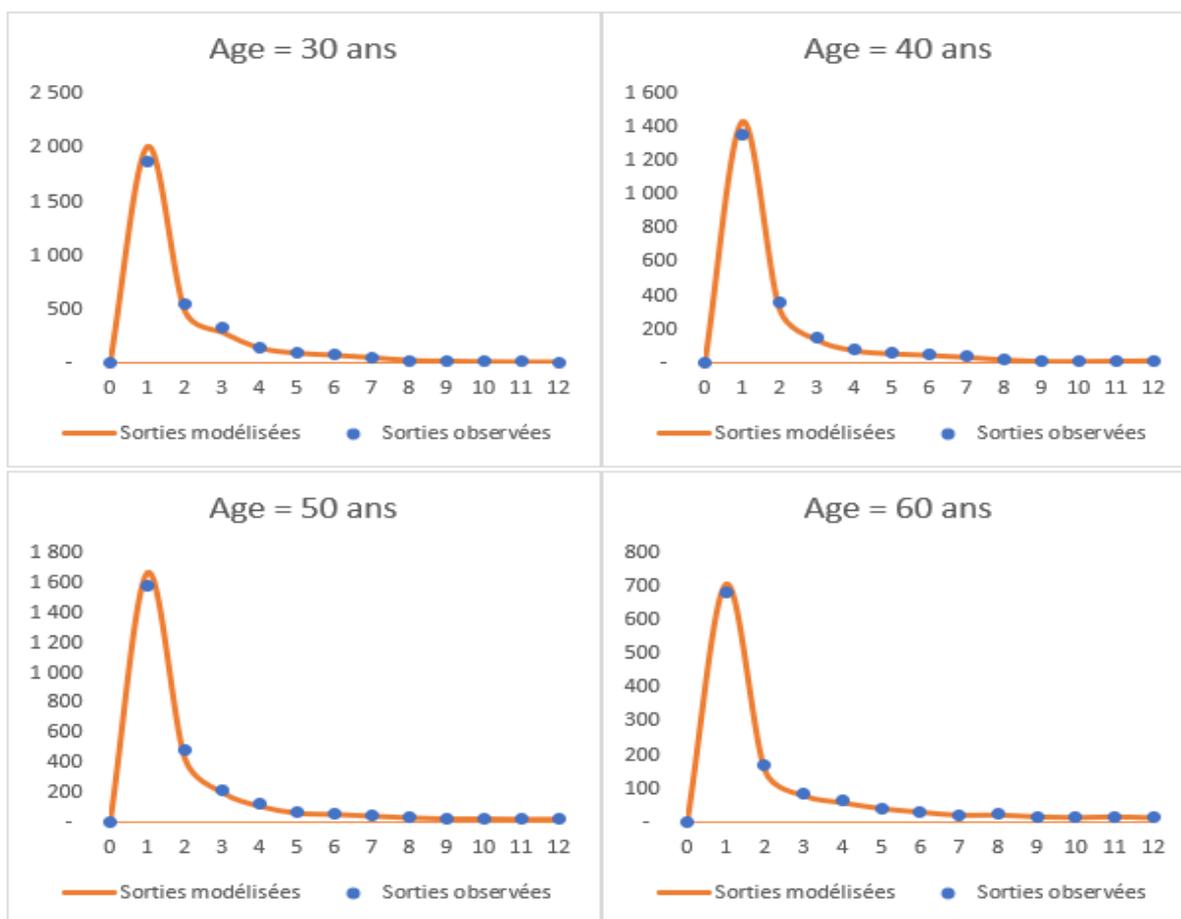


Figure 29- Sorties modélisées vs. Sorties observées : zoom par âge

Nous faisons le même constat sur la bonne adéquation du modèle à nos données pour les différents âges représentés. Par ailleurs, une légère surestimation des sorties est constatée pour le premier mois d'ancienneté. Elle est compensée par une légère sous-estimation pour les mois suivants.

4. Détermination de la loi de maintien en incapacité

Nous pouvons à présent calculer les L_x de la loi d'expérience à partir des taux de sortie lissés en procédant comme suit :

$$L_{x,0} = 10\,000$$

$$L_{x,1} = (1 - q_{x,1}^{\text{lissé}}) * L_{x,0}$$

$$L_{x,2} = (1 - q_{x,2}^{\text{lissé}}) * L_{x,1}$$

$$\dots$$

$$L_{x,12} = (1 - q_{x,12}^{\text{lissé}}) * L_{x,11}$$

Pour rappel, $L_{x,k}$ représente le nombre d'individus entrés en incapacité à l'âge x et toujours dans cet état dans k mois.

Ci-dessous un extrait de la loi obtenue :

		Ancienneté dans l'état d'incapacité (en mois)					
		0	1	2	3	...	12
Age à l'entrée en incapacité	21 ou -	10000	2885	1517	864	...	71

	30	10000	3775	2293	1418	...	156

	50	10000	4079	2572	1870	...	567

60	10000	4376	3156	2531	...	738	

Figure 30- Extrait de la loi d'expérience

Comme nous l'avons vu précédemment, nous avons modélisé les taux de sortie d'incapacité jusqu'à 12 mois d'ancienneté et pour un âge d'entrée en incapacité limité à 60 ans par soucis de fiabilité des résultats. Pour étendre la loi d'expérience jusqu'à 36 mois d'ancienneté et au-delà de 60 ans, nous effectuons un raccordement avec une loi de référence.

Chapitre 6 : Prolongement de la loi d'expérience

Nous avons choisi d'utiliser la table BCAC 2010 comme référence pour effectuer le raccordement avec notre loi d'expérience. En effet, cette table est toujours en vigueur alors que la table BCAC 2013 n'a pas encore été homologuée depuis sa construction. De plus, la loi de maintien de 2010 est plus proche

des observations de notre portefeuille que celle de 2013. En effet, selon une étude³² réalisée en 2014 par le cabinet Prim'Act, la table de 2013 surestime la durée de maintien en incapacité par rapport à celle de 2010 pour les anciennetés faibles et sous-estime légèrement cette durée pour les anciennetés élevées.

D'un point de vue pratique, un double prolongement est effectué :

- Prolongement horizontal pour $x \in [61 ; 66]$:

$$L_{x,t}^{Exp} = \frac{L_{x,t}^{BCAC}}{L_{x-1,t}^{BCAC}} * L_{x-1,t}^{Exp} \quad \forall t \in [1 ; 12]$$

- Prolongement vertical pour $t \in [13 ; 36]$:

$$L_{x,t}^{Exp} = \frac{L_{x,t}^{BCAC}}{L_{x,t-1}^{BCAC}} * L_{x,t-1}^{Exp} \quad \forall x \in [20 ; 66]$$

Ci-dessous une illustration du prolongement effectué :

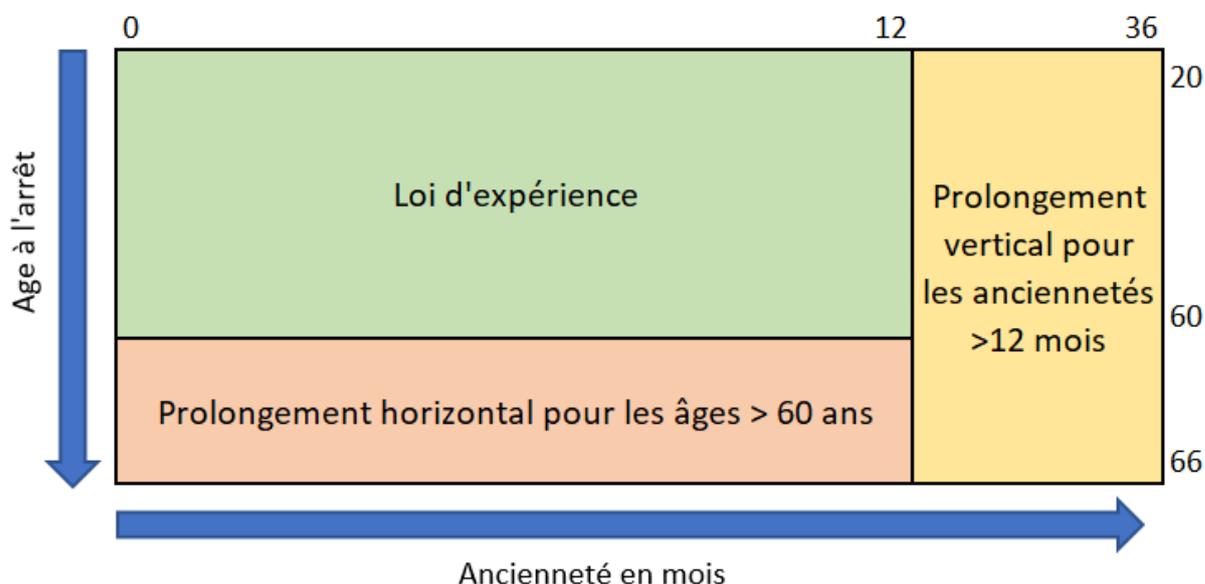


Figure 31- Illustration du prolongement

Nous obtenons ainsi une loi de maintien sur 36 mois avec un âge d'entrée en incapacité allant jusqu'à 66 ans.

Pour valider le bon fonctionnement du raccordement, nous regardons l'allure de la loi d'expérience pour différents âges :

³² Cabinet Prim'Act, A. Kamega, F. Planchet, R. Wolfrum (2014) : Présentation et comparaison des nouvelles tables BCAC.

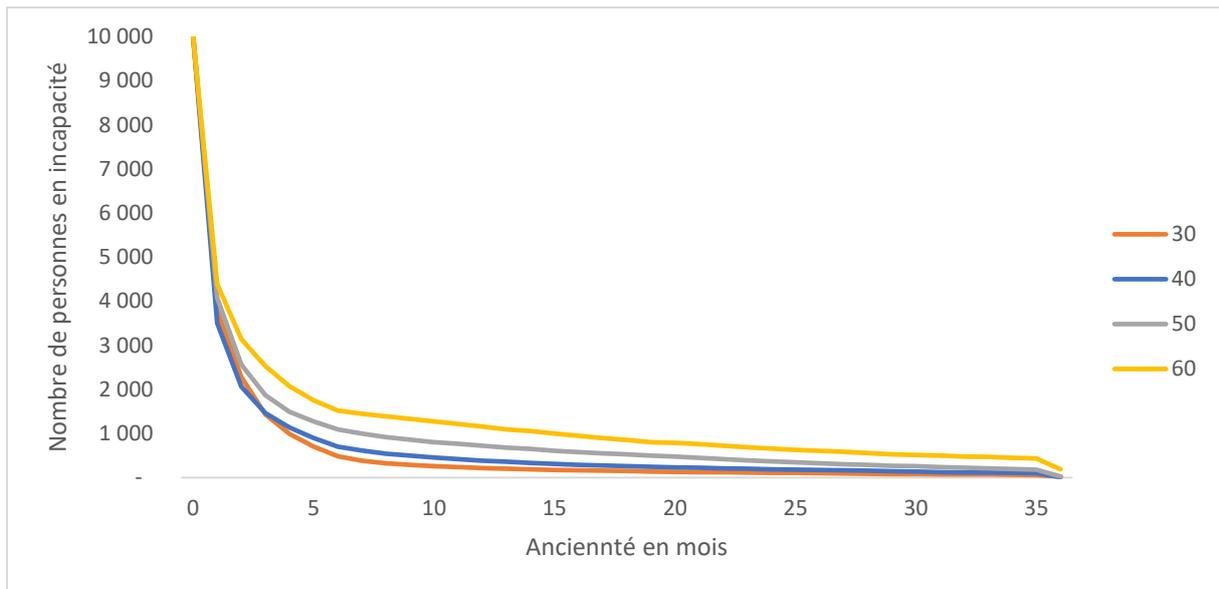


Figure 32- Loi de maintien en incapacité : zoom par âge

Les allures des courbes sont cohérentes en fonction de l'âge. Nous observons bien que la probabilité de maintien en incapacité augmente avec l'âge. Cela confirme encore une fois que l'âge est un facteur aggravant du risque de maintien en incapacité.

Nous observons également que la loi de maintien décroît en fonction du temps avec une pente très forte lors des premiers mois d'ancienneté. Ce résultat est expliqué par le fait qu'un assuré a plus de chance de sortir de l'incapacité au début de son arrêt que vers la fin.

Aucune variation atypique n'est observée graphiquement, en particulier avant et après 12 mois d'ancienneté.

Partie 4 : Mise à l'épreuve de la loi d'expérience

Dans cette partie, la loi d'expérience sera comparée à la loi du BCAC.

Les provisions mathématiques de maintien en incapacité seront calculées en utilisant les deux lois. Les résultats seront ensuite, confrontés pour étudier l'impact de l'utilisation de la nouvelle table.

Une adaptation de la règle actuelle de sélection des encours sera proposée dans le but d'avoir une évaluation réaliste du risque.

Chapitre 1 : Comparaison avec la table règlementaire

Nous comparons le nombre de personnes en état d'incapacité au début de chaque mois d'ancienneté issu de la loi règlementaire avec celui obtenu à travers l'utilisation de la loi d'expérience :

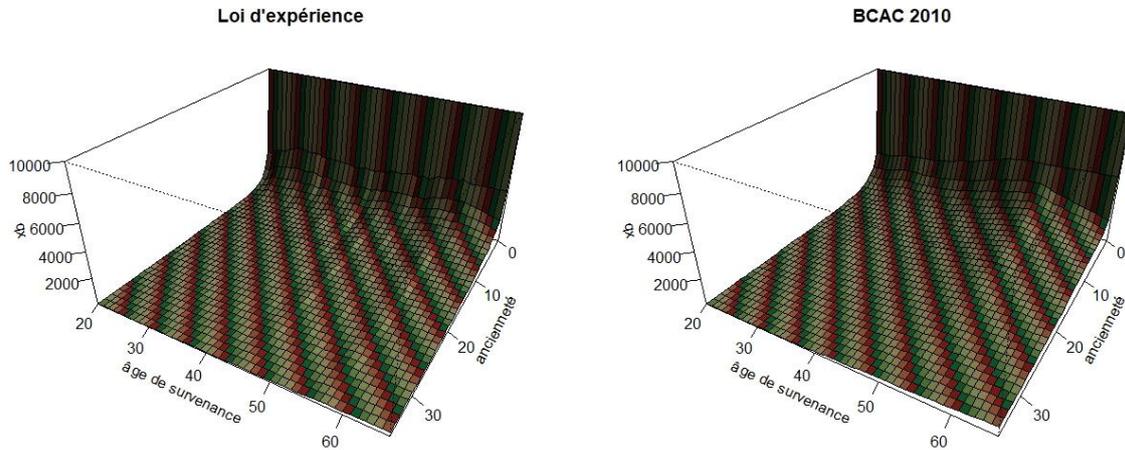


Figure 33- Comparaison des lois de maintien en incapacité

La loi d'expérience présente une allure très proche de celle du BCAC avec une surface plus régulière.

Pour mieux visualiser le positionnement de la loi d'expérience par rapport à la loi du BCAC, nous représentons graphiquement les deux lois pour les âges de 30, 40, 50 et 60 ans :

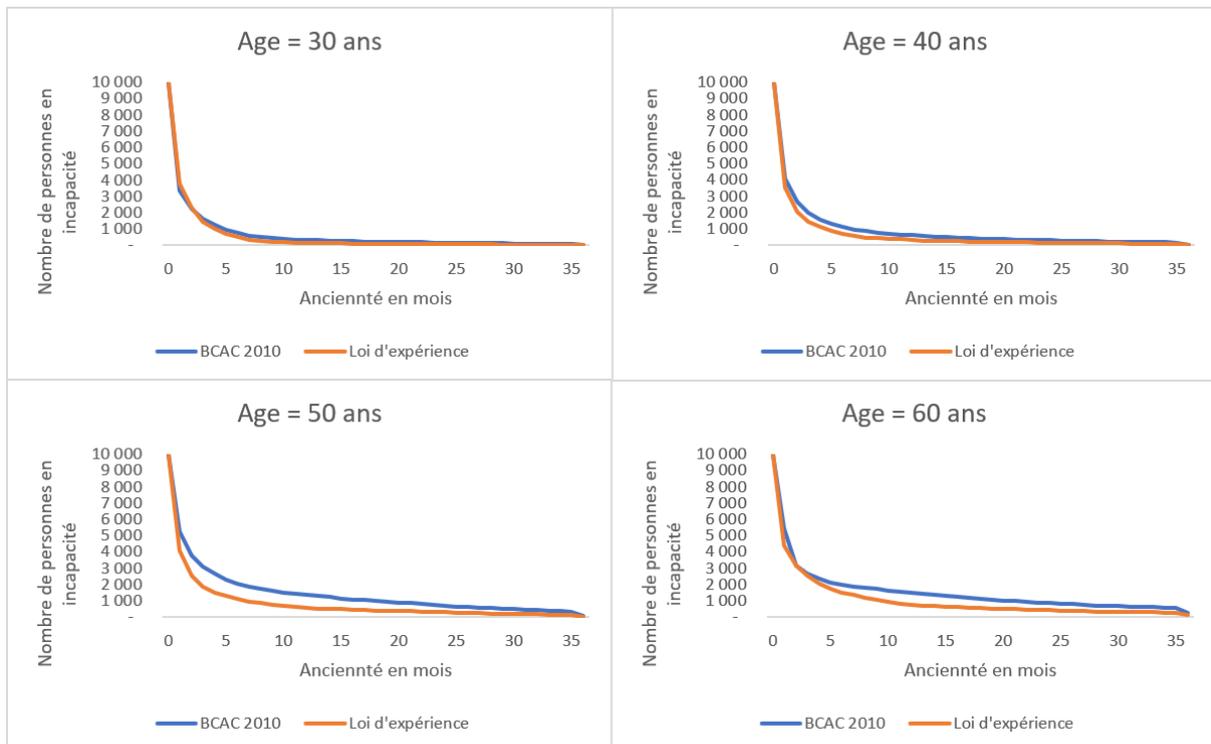


Figure 34- Comparaison des lois de maintien en incapacité : zoom par âge

Nous remarquons que la courbe représentative de la loi d'expérience est quasiment toujours en dessous de celle de la loi réglementaire quel que soit l'âge d'entrée en incapacité. Ceci montre que la table du BCAC surestime la durée de maintien en incapacité par rapport à la table d'expérience.

Nous constatons aussi que la pente de la loi modélisée est plus forte que celle du BCAC pour les anciennetés faibles, ce qui signifie que les assurés sortent plus rapidement de l'état d'incapacité. En revanche, les pentes des deux lois sont similaires au-delà de 12 mois d'ancienneté à cause du raccordement.

Les écarts entre les deux lois sont plus marqués pour l'âge d'entrée en incapacité de 50 ans. Dans ce qui suit, nous faisons un zoom sur les 12 premiers mois d'ancienneté en représentant le taux de sortie d'incapacité³³ mois par mois pour cet âge :

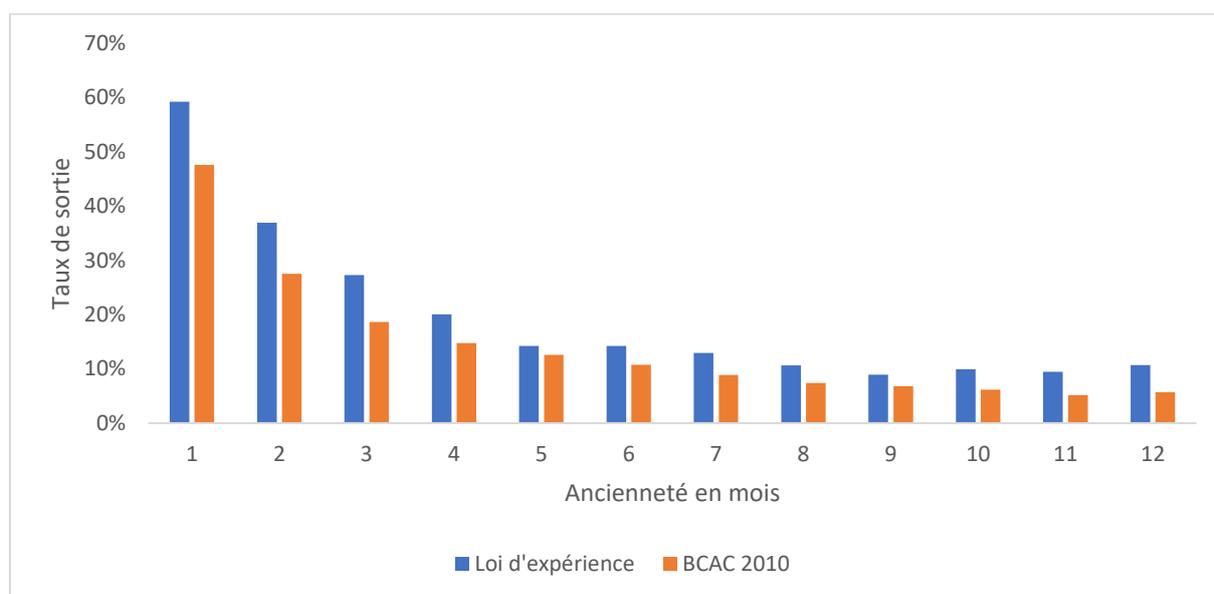


Figure 35- Comparaison des taux de sortie pour un âge d'entrée en incapacité de 50 ans

Le principal écart sur les taux de sortie se situe dans le premier mois d'ancienneté : 59% des arrêts sont clos à la fin du premier mois selon la loi d'expérience contre seulement 48% selon la loi du BCAC. Ce résultat est cohérent avec la spécificité de la franchise courte du portefeuille étudié.

L'écart reste sensiblement élevé lors du deuxième et du troisième mois d'ancienneté avec un delta de 9 points entre les taux de sortie donnés par les deux lois. Il descend ensuite sous la barre de 5 points à partir du quatrième mois.

A la fin de la première année d'ancienneté, 94% des arrêts sont clos selon la loi d'expérience contre 86% selon la loi du BCAC.

³³ Taux de sortie au mois M= nombre de sorties à la fin du mois M / nombre de personnes en incapacité au début du mois M.

Chapitre 2 : Utilisation de la loi d'expérience

Nous allons à présent calculer les provisions mathématiques de maintien en incapacité et étudier l'impact que représente l'utilisation de la loi d'expérience par rapport à la loi règlementaire.

1. Calcul des provisions de maintien en incapacité

Nous rappelons que pour un individu entré en incapacité à l'âge x et encore dans cet état t mois plus tard, le montant de la provision mathématique à constituer pour le versement d'une prestation de 1€ est défini comme suit :

$$PM_{x,t}^{INC} = \frac{1}{2 L_{x,t}^{INC}} \sum_{k=t}^{35} \frac{L_{x,k}^{INC}}{(1+i)^{\frac{k-t}{12}}} + \frac{L_{x,k+1}^{INC}}{(1+i)^{\frac{k+1-t}{12}}}$$

Avec,

- i : taux d'actualisation annuel.
- $L_{x,k}^{INC}$: nombre d'individus entrés en incapacité à l'âge x et toujours dans cet état dans k mois.

A l'aide de cette formule, nous pouvons calculer les provisions unitaires (coefficients) pour chaque âge et pour chaque ancienneté donnée. Nous obtenons ainsi une table de provisionnement, de même dimension que la table de maintien en incapacité, appelée barème de provisionnement.

Ci-dessous un extrait du barème obtenu pour un taux d'actualisation nul :

		Ancienneté dans l'état d'incapacité (en mois)					
		1	2	3	4	...	36
Age à l'entrée en incapacité	20	1,41	1,88	2,54	3,41	...	0,65

	30	1,99	2,48	3,09	3,84	...	0,65

	50	2,98	5,11	6,87	8,27	...	0,58

66	3,33	6,48	9,00	10,88	...	0,77	

Figure 36- Extrait du barème de provisionnement

A partir du barème de provisionnement, nous pouvons construire les courbes d'évolution de la provision unitaire en fonction de l'âge et de l'ancienneté dans l'arrêt :

Barème de provisionnement

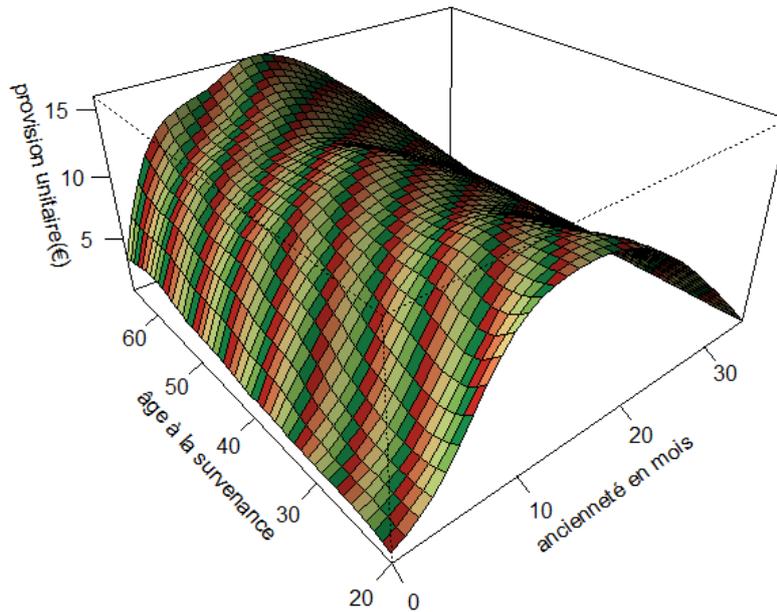


Figure 37- Evolution de la provision unitaire en fonction de l'âge et de l'ancienneté dans l'arrêt

Globalement, l'allure de la courbe est la même quel que soit l'âge d'entrée en incapacité. En effet, la provision augmente pendant les premiers mois de l'arrêt pour arriver à un pic en fonction de l'âge entre 13 mois et 15 mois, puis elle se met à décroître pour s'annuler au bout de 36 mois d'ancienneté.

Ci-dessous un zoom sur les âges de 30, 40, 50 et 60 ans :

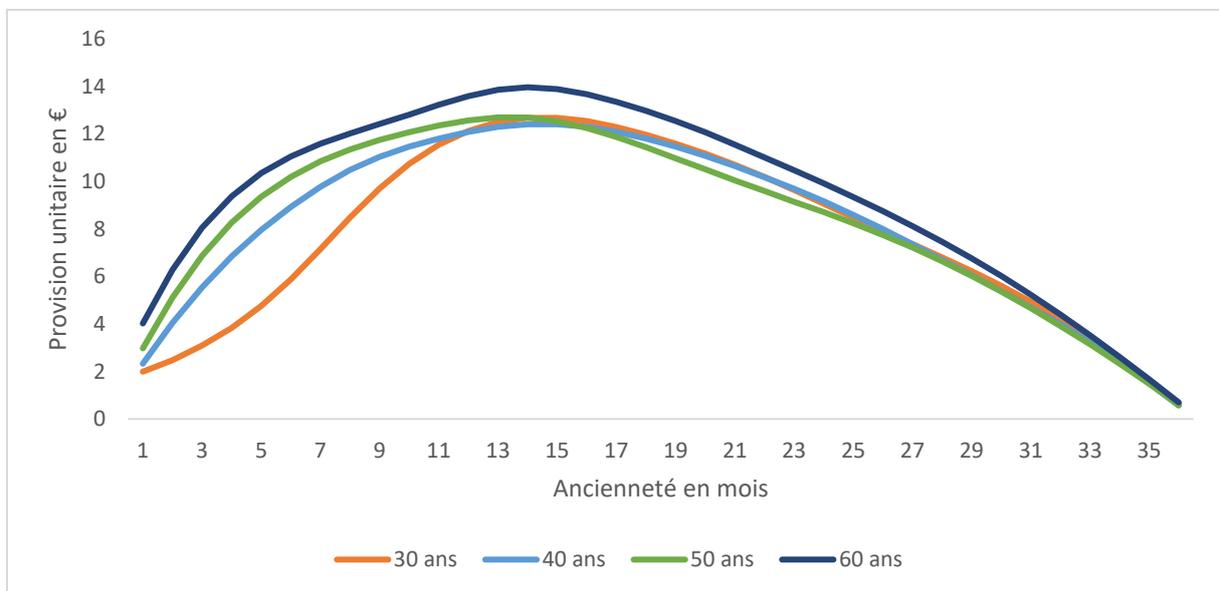
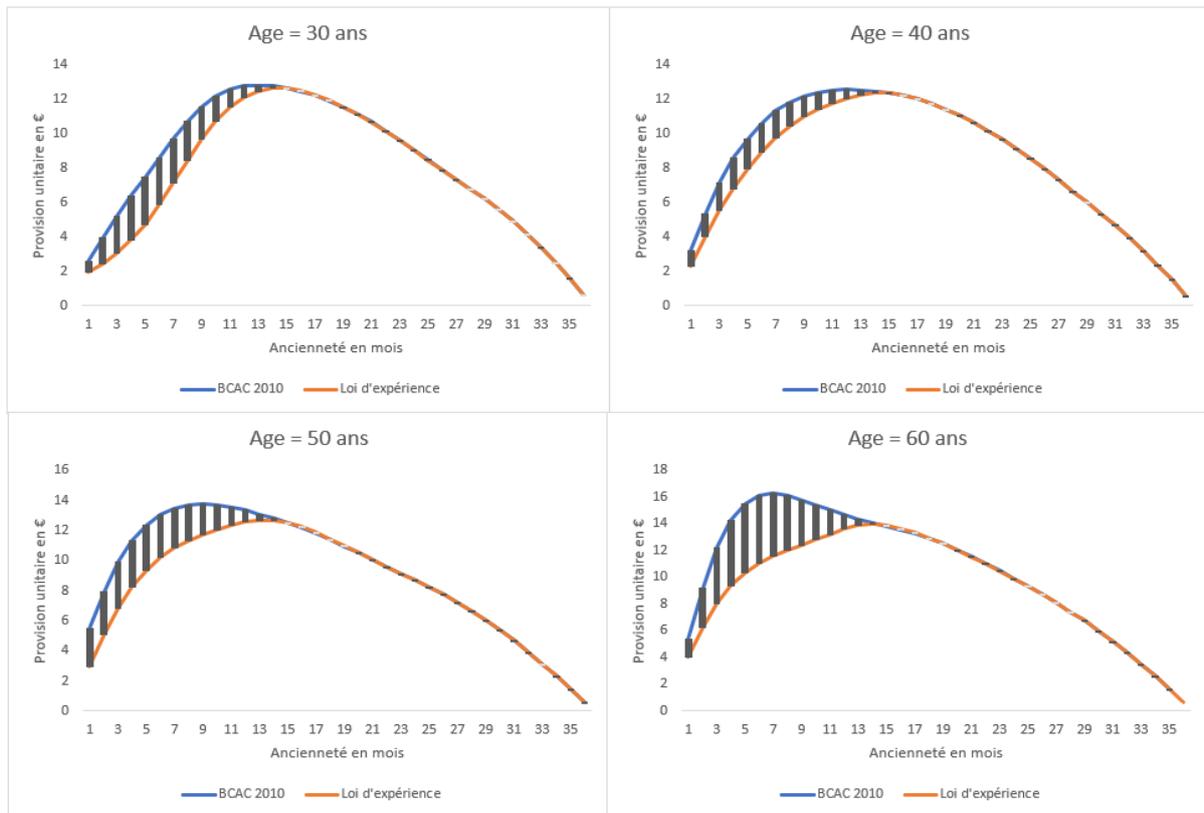


Figure 38- Evolution de la provision unitaire en fonction de l'ancienneté en arrêt : zoom par âge

Nous remarquons que le niveau de la provision est plus élevé avec l'âge d'entrée en incapacité au début de la courbe. Cependant, l'écart à tendance à diminuer à partir de 12 mois d'ancienneté. Autrement dit, l'influence de l'âge dans l'explication du maintien en incapacité diminue après une année d'ancienneté dans l'état.

Après avoir construit le barème de provisionnement de la loi d'expérience, il est intéressant de le comparer à celui de la loi du BCAC pour les âges représentés précédemment :



Ce graphique met en évidence que les provisions calculées avec la loi d'expérience sont systématiquement moins élevées que celles calculées avec la loi du BCAC pendant les 12 premiers mois. Les courbes se superposent ensuite par construction.

Nous constatons également que l'aire séparant les deux courbes est plus grande pour les âges d'entrée en incapacité les plus élevés.

Remarque : Pour déterminer le coefficient de provisionnement pour un âge et une ancienneté exacts, nous consultons 4 valeurs du barème à chaque fois.

Notons :

- *Age* : l'âge à la date d'arrêt de travail (la durée exacte en année séparant les dates d'arrêt de travail et de naissance).
- *Age 1* : la valeur entière inférieure de *Age*.
- *Age 2* : la valeur entière supérieure de *Age*.
- *Anc* : l'ancienneté de l'arrêt de travail (la durée exacte en mois séparant les dates d'évaluation des PM et d'arrêt de travail).
- *Anc 1* : la valeur entière inférieure de *Anc*.
- *Anc 2* : la valeur entière supérieure de *Anc*.

Quatre valeurs du barèmes sont consultées :

- $M1 = Coefficient(Age\ 1, Anc\ 1)$
- $M2 = Coefficient(Age\ 2, Anc\ 1)$

- $M3 = Coefficient(Age\ 1, Anc\ 2)$
- $M4 = Coefficient(Age\ 2, Anc\ 2)$

Une interpolation est effectuée pour retenir le coefficient pour le calcul de la provision mathématique :

$$Coefficient(Age, Anc) = (Anc\ 2 - Anc) * [(Age\ 2 - Age) * M1 + (Age - Age1) * M2] + (Anc - Anc1) * [(Age\ 2 - Age) * M3 + (Age - Age\ 1) * M4]$$

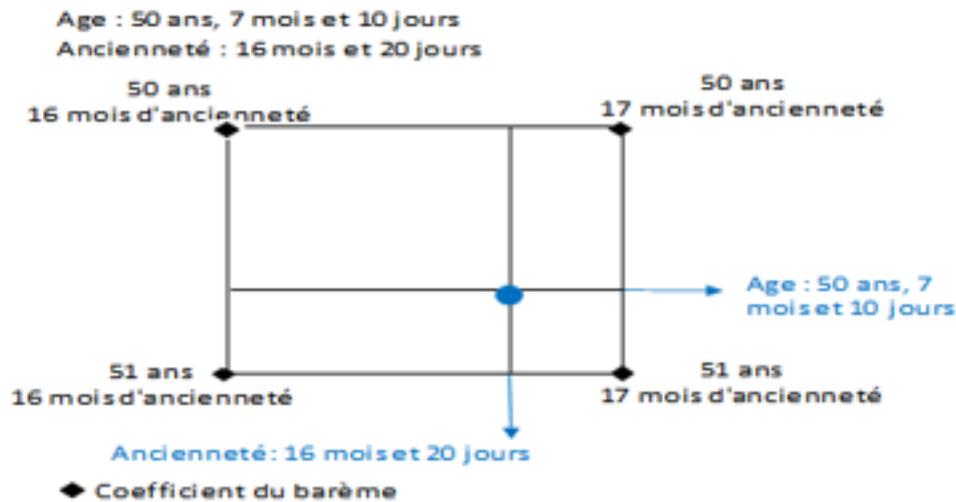


Figure 40- Illustration de la méthode de détermination du coefficient de provisionnement

Le calcul des provisions mathématiques repose sur le produit des prestations atteintes en date d'inventaire avec le coefficient correspondant au barème à un âge et une ancienneté donnés.

2. Etude d'impact sur le provisionnement

La méthodologie suivante a été suivie pour l'étude d'impact sur le provisionnement :

1. Détermination de la liste des encours incapacité à fin 2019 et à fin 2020 pour les sinistres survenus entre le 1^{er} janvier 2019 et le 31 décembre 2019 ;
2. Calcul des PM à fin 2019 et à fin 2020 (en utilisant les deux lois de maintien en incapacité) ;
3. Comparaison de la variation des PM avec le montant réel des prestations payées en 2020 pour les sinistres ayant fait l'objet d'une indemnisation au titre de la période 01/01/2020 – 31/12/2020.

Remarques :

- Les prestations concernent les sinistres provisionnés à fin 2019 mais aussi les sinistres tardifs (indemnisés pour la première fois en 2020). En raison des boni importants, aucune provision pour sinistre inconnu n'est constituée dans les comptes de résultat relevant de ce périmètre.
- Les données de la survenance 2019 n'ont pas servi à la construction de la loi de maintien. Nous accordons donc une attention particulière à cette survenance pour tester la robustesse de la modélisation.

Survenance	Nombre d'arrêts	Age moyen à l'arrêt	Durée moyenne de l'arrêt (en jours)
2019	14 685	42	64
2018	13 518	42	64

Tableau 9- Comparaison entre la survenance 2019 et la survenance 2018 à date d'arrêt identique

A travers les statistiques présentées dans le tableau précédent, nous vérifions que la survenance 2019 présente les mêmes caractéristiques que la survenance 2018 (qui fait partie de la base d'étude) en termes d'âge moyen et de durée moyenne de l'arrêt (données observées au 31/12/N+1).

Avant de passer au calcul des provisions mathématiques, nous précisons les paramètres techniques retenus³⁴ en ce qui concerne le taux d'actualisation et la règle de sélection des encours :

Inventaire	Taux d'actualisation	Règle de sélection des encours
31/12/2019	0%	Dernière période indemnisée <= 4 mois par rapport à la date d'inventaire
31/12/2020	0%	Dernière période indemnisée <= 4 mois par rapport à la date d'inventaire

Tableau 10- Paramètres techniques pour le calcul des provisions

Les résultats des calculs de PM sont présentés dans le tableau ci-après :

Inventaire	Nombre encours	PM Incapacité		Ecart en €	Ecart en %
		BCAC	Expérience		
2019	2 717	20 644 939 €	15 766 198 €	-4 878 741 €	-24%
2020	555	4 200 048 €	4 193 523 €	-6 526 €	0%

Tableau 11- PM incapacité en cours au 31/12/N

A l'inventaire 2019, le montant des provisions mathématiques a diminué de 24% en utilisant la loi d'expérience. Cette diminution a pour origine la différence d'évolution des lois de maintien. En effet, la provision mathématique dépend principalement de la pente de la courbe de maintien qui décroît plus vite avec les données propres à notre portefeuille.

A l'inventaire 2020, nous avons un provisionnement quasi-identique par construction de la loi d'expérience. En effet, le changement de la loi de maintien n'a pratiquement pas d'impact sur le calcul des PM puisque la durée de maintien en incapacité dépasse 12 mois pour les sinistres survenus au plus tard à fin 2019. Le faible écart de 6,5 K€ s'explique par la méthode de calcul de la provision mathématique (moyenne d'un flux à terme échu et d'un flux à terme d'avance) ainsi que l'interpolation effectuée pour retenir le coefficient de provisionnement à partir du barème³⁵.

Afin de tester la fidélité de la loi modélisée à notre portefeuille, nous faisons une analyse boni/mali en comparant la variation des PM entre deux inventaires aux prestations réellement payées :

³⁴ Conformément aux règles de calcul des provisions mathématiques adoptées par le groupe AG2R La Mondiale.

³⁵ Concerne les anciennetés (en jours) de l'arrêt de travail entre 12 mois et 13 mois.

	BCAC	Expérience
Prestations 2020	6 327 270 €	
PM incapacité 2019	20 644 939 €	15 766 198 €
PM incapacité 2020	4 200 048 €	4 193 523 €
PM incapacité (variation)	16 444 891 €	11 572 675 €
Boni (+) / Mali (-)³⁶	10 117 621 €	5 245 405 €
Ecart Boni / Mali en €	-4 872 216 €	
Ecart Boni / Mali en %	-48%	
Estimation / Réel³⁷	260%	183%
Boni / PM 2019	49%	33%

Tableau 12- Analyse boni/mali BCAC vs. Expérience

L'utilisation de la loi d'expérience permet de réduire le boni observé en utilisant la table du BCAC de 48%. En effet, La table du BCAC surestime la charge réelle de 160% alors que la loi d'expérience surestime la charge réelle de 83%.

Chapitre 3 : Adaptation de la règle de sélection des encours

Bien que l'utilisation de la nouvelle loi permette de réduire considérablement le sur-provisionnement généré par l'utilisation de la loi réglementaire, l'écart entre la provision calculée et les prestations réelles reste sensiblement élevé.

En réalité, ce constat est directement lié à la règle actuelle de sélection des encours : sur les 2 717 encours provisionnés à fin 2019, seulement 890 ont continué à avoir des prestations en 2020 (soit 33%).

Cette règle conduit systématiquement à une surestimation fictive des montants à provisionner. Il est donc nécessaire de neutraliser ce biais avant de déterminer l'impact de l'utilisation de la loi d'expérience.

En général, nous disposons d'un recul d'au moins 3 mois entre la date d'inventaire et la date d'arrêté des flux financiers dans le cadre de la production des comptes clients (pour avoir une année complète de cotisations notamment).

Il est possible d'exploiter ce recul par rapport à la date d'inventaire pour limiter au maximum le nombre de dossiers présents dans la liste des encours et qui sont déjà clos.

Une nouvelle sélection des encours au 31/03/N+1³⁸ sera proposée dans ce qui suit. L'objectif est double :

- Réduire le sur-provisionnement en tenant compte des sorties avérées en fin d'année ;

³⁶ Ecart entre la variation des PM incapacité et les prestations 2020.

³⁷ Estimation / Réel = PM incapacité (variation) / Prestations 2020.

³⁸ Hors survenance N+1.

- Anticiper les éventuels sinistres tardifs (surtout en l'absence de provisions pour sinistres inconnus dans les comptes client).

Evaluons d'abord l'impact que représente le changement de la règle de sélection des encours sur le calcul des provisions mathématiques à l'inventaire 2019 :

Sélection des encours	Nombre encours	PM Incapacité	
		BCAC	Expérience
31/12/2019	2 717	20 644 939 €	15 766 198 €
31/03/2020	2 209	13 396 592 €	10 057 957 €
Variation	-19%	-35%	-36%

Tableau 13- Impact du changement de la règle de sélection des encours sur le calcul des PM au 31/12/2019

Il est clair que la prise en compte des encours au 31/03/2020 a un impact significatif sur le provisionnement : -19% sur le nombre de dossiers provisionnés et -35% à minima pour les PM incapacité en cours selon la loi utilisée.

Les résultats des calculs de PM (après la prise en compte de la nouvelle règle de sélection des encours) sont présentés dans le tableau ci-après :

Inventaire	Nombre encours	PM Incapacité		Ecart en €	Ecart en %
		BCAC	Expérience		
2019	2 209	13 396 592 €	10 057 957 €	-3 338 635 €	-25%
2020	444	3 235 807 €	3 230 722 €	-5 085 €	0%

Tableau 14- PM incapacité en cours au 31/12/N (sélection des encours au 31/03/N+1)

A l'inventaire 2019, le montant des provisions mathématiques a diminué de 25% en utilisant la loi d'expérience. Nous retrouvons pratiquement les mêmes évolutions que précédemment lors du calcul avec la liste des encours au 31/12/2019.

Le changement de la loi de maintien n'a pas d'impact sur le calcul des PM à fin 2020 pour les mêmes raisons que précédemment.

Reprenons désormais notre analyse boni/mali :

	BCAC	Expérience
Prestations 2020	6 327 270 €	
PM incapacité 2019	13 396 592 €	10 057 957 €
PM incapacité 2020	3 235 807 €	3 230 722 €
PM incapacité (variation)	10 160 786 €	6 827 235 €
Boni (+) / Mali (-)	3 833 515 €	499 965 €
Ecart Boni / Mali en €	-3 333 550 €	
Ecart Boni / Mali en %	-87%	
Estimation / Réel	161%	108%
Boni / PM 2019	29%	5%

Tableau 15- Analyse boni/mali BCAC vs. Expérience (sélection des encours au 31/03/N+1)

Avec une sélection des encours au 31/03/N+1, l'utilisation de la loi d'expérience permet de réduire le boni observé en utilisant la table du BCAC de 87% (vs. 48% avec la règle en vigueur de sélection des encours).

La table du BCAC surestime la charge réelle de 61% (vs. 160% avec la règle en vigueur de sélection des encours) alors que la loi d'expérience surestime la charge réelle de 8% seulement (vs. 83% avec la règle en vigueur de sélection des encours).

Au regard de ces résultats, nous pouvons conclure que la loi d'expérience est plus fidèle à la réalité du portefeuille étudié. En effet, le provisionnement obtenu avec cette loi est plus en adéquation avec la charge de prestation réelle, contrairement à la table règlementaire qui génère des bonis significatifs.

Il est à noter qu'un « simple » changement de la règle de sélection des encours n'est pas suffisant. En effet, un sur-provisionnement de l'ordre de 61% est mis en évidence lors de l'utilisation de la table du BCAC. De même, utiliser la loi d'expérience en maintenant la règle de sélection actuelle conduit à un sur-provisionnement de l'ordre de 83%. La solution que nous proposons consiste à utiliser la loi d'expérience avec un décalage de la date d'observation des sinistres de 3 mois. Cette solution permet d'atteindre l'objectif attendu, à savoir fournir une estimation réaliste du risque.

Conclusion

Au terme de ce travail, nous avons construit une loi de maintien en incapacité sur un portefeuille de contrats caractérisés par une franchise courte. L'objectif était d'assurer un meilleur suivi des régimes de branches concernés par cette spécificité.

Après une étape cruciale de fiabilisation des données, une analyse statistique du portefeuille a été menée. Elle a permis de mettre en exergue une certaine hétérogénéité de la population. De plus, l'application de la méthode des forêts aléatoires a montré que le salaire, l'âge, la franchise et le type d'arrêt sont des paramètres discriminants du risque de maintien en incapacité. Compte tenu de la composition de nos données, l'âge a finalement été retenu comme unique variable de segmentation.

Pour modéliser la loi d'expérience, nous avons opté pour des méthodes non paramétriques. Ce choix est justifié par la non connaissance à priori de la loi théorique qui pourrait être utilisée. L'estimateur de Kaplan-Meier a été ainsi retenu pour la construction des taux bruts de sortie d'incapacité. Par la suite, un lissage de ces taux par la méthode de Whittaker Henderson a été effectué pour diminuer les irrégularités. Une fois l'ajustement validé statistiquement et graphiquement, nous avons pu vérifier la bonne adéquation du modèle aux données du portefeuille.

En raison de la volumétrie des données, nous avons modélisé les taux de sortie d'incapacité jusqu'à 12 mois d'ancienneté et pour un âge d'entrée en incapacité limité à 60 ans. Pour étendre la loi d'expérience jusqu'à 36 mois d'ancienneté et au-delà de 60 ans, un raccordement a été effectué avec la table règlementaire.

Pour finir, nous avons calculé les provisions mathématiques de maintien en incapacité et nous avons évalué l'impact que représente l'utilisation de la nouvelle loi plutôt que la loi du BCAC. Il en ressort que la loi d'expérience est plus fidèle à la réalité du portefeuille étudié. En effet, le provisionnement obtenu avec cette table est plus en adéquation avec la charge de prestation réelle.

Toutefois, un suivi permanent de la loi modélisée est nécessaire afin de s'assurer de son adéquation à la réalité du portefeuille. En effet, tout changement dans la structure de ce dernier devrait être accompagné d'une réadaptation de la loi.

En définitive, il est indispensable de signaler que les résultats de l'utilisation de la loi d'expérience dépendent fortement de la règle de sélection des encours. A ce titre, une adaptation de cette règle a été proposée dans le but d'avoir une évaluation réaliste du risque tout en respectant les contraintes opérationnelles de production des comptes clients.

Pour aller plus loin, ce travail pourrait être prolongé en étudiant la probabilité de passage en invalidité afin d'avoir une vision complète du risque arrêt de travail.

Bibliographie

AIT EL HOUSS R. (2018) *Maîtrise du risque incapacité appliquée à un portefeuille de professionnels de santé*. Mémoire d'actuariat, ENSAE.

Centre Technique des Institutions de Prévoyance (2018) *Etat des lieux des accords de branches en prévoyance*. Communiqué de presse du 21 mars 2018.

CHESNEAU C. (2020) *Introduction aux arbres de décision (de type CART)*. Support de cours 2020-2021, Université de Caen.

FETOUI N. (2015) *Impact de l'utilisation des tables d'expérience sur le provisionnement en prévoyance*. Mémoire d'actuariat, Université de Paris Dauphine.

GABRIELLE J. (2019) *Dépendance totale : Modélisation du risque de longévité pour un portefeuille de rentes individuelles*. Mémoire d'actuariat, ISFA.

KAMEGA A., PLANCHET F., WOLFRUM R. (2014) *Présentation et comparaison des nouvelles tables BCAC*. Support de présentation, Cabinet Prim'Act.

LEBAMBA-MOUSSA R. (2017) *Construction d'indicateurs et des lois de maintien en incapacité afin de mesurer l'hétérogénéité : application sur un portefeuille de prévoyance collective*. Mémoire d'actuariat, CNAM.

LE GOFF G. (2019) *Actuariat des régimes de protection sociale*. Support de cours 2018-2019, CNAM.

MORLON P. (2017) *Construction d'une loi de changement de catégorie d'invalidité et étude d'impact sur les provisions mathématiques*. Mémoire d'actuariat, Université de Strasbourg.

NGUYEN K. (2018) *Méthodes de provisionnement du maintien en incapacité des contrats dits franchises courtes*. Mémoire d'actuariat, CEA.

PLANCHET F., THEROND P. (2006) *Modèles de Durée - Applications actuarielles*, Economica.

PLANCHET F. (2020) *Modèles de durée*. Support de cours 2020-2021, ISFA.

ROUSSEL J., ZAOUÏ R. (2014) *Loi expérimentale de maintien en incapacité temporaire avec un pas quotidien*. Mémoire d'actuariat, CEA.

Note de synthèse

Contexte et problématique

Pour pouvoir honorer leurs engagements vis-à-vis de leurs assurés, les assureurs doivent constituer des provisions représentatives des risques auxquels ils sont exposés.

En prévoyance, l'estimation des provisions pour sinistres en cours représente un exercice crucial pour les organismes assureurs. Ces provisions, constituées afin de faire face aux versements futurs des prestations périodiques, font appel au calcul de probabilités ainsi qu'à l'actualisation des flux financiers.

Dans le cadre de la production des comptes clients, le provisionnement du risque arrêt de travail est actuellement effectué à l'aide des tables réglementaires établies par le BCAC. Ces dernières sont basées sur les portefeuilles des principales compagnies d'assurance françaises.

Cependant, leur utilisation peut conduire à surestimer les provisions mathématiques dans certaines conditions. C'est le cas des contrats de prévoyance collective qui prévoient une garantie incapacité avec une franchise courte. Sur ce périmètre, des bonis récurrents et significativement élevés sont remarqués notamment pour le risque de maintien en incapacité.

Le portefeuille étudié dans ce mémoire concerne des contrats avec une franchise continue dont le délai ne dépasse pas 10 jours. Il s'agit d'une particularité présente dans plusieurs contrats collectifs assurés par AG2R La Mondiale. Cette spécificité est liée à des obligations de certains régimes de branches.

Lors de l'établissement des comptes de résultats, un ratio S/P largement déficitaire est remarqué systématiquement sur la dernière survenance présentée. Ce ratio ne traduit pas la réalité de la situation et par conséquent, il ne permet pas de tirer des conclusions sur un éventuel déséquilibre technique du risque incapacité et d'une façon plus générale du régime assuré.

Dans cette optique s'inscrit la volonté de construire une loi de maintien en incapacité qui reflète mieux la sinistralité de ce portefeuille. L'objectif est d'évaluer plus précisément les provisions correspondantes et par conséquent assurer un meilleur pilotage des régimes concernés.

Données utilisées

Le portefeuille est constitué de contrats collectifs en gestion directe issus de différentes Conventions Collectives Nationales. La base d'étude initiale comporte 107 494 observations. Chaque observation représente un arrêt de travail survenu entre le 15/10/1998 et le 31/12/2018. Les données sont observées au 31/12/2019. L'observation des données après un recul d'un an par rapport à la dernière date de survenance permet d'éviter les problématiques de déclarations tardives et d'avoir des données majoritairement complètes. Par ailleurs, afin de tester la robustesse de la modélisation, les calculs et les comparaisons des provisions mathématiques ont été effectués sur les sinistres de la survenance 2019.

Après une étape cruciale de fiabilisation des données, une analyse statistique du portefeuille a été menée. Il en ressortait que la population étudiée est majoritairement féminine. Elle est essentiellement constituée de non cadres, mariés et exerçant dans le secteur de la santé. La principale

cause d'entrée en incapacité est la maladie. L'âge d'entrée en incapacité est en moyenne autour de 41 ans. Les sinistres sont clos dès la première année dans 97% des cas.

Les résultats statistiques ont également montré une certaine hétérogénéité de la population au regard de la durée de l'arrêt. Pour aller plus loin dans la segmentation des profils de risques, la méthode des forêts aléatoires a été appliquée. Elle a mis en exergue que le salaire, l'âge, la franchise et le type d'arrêt sont des paramètres discriminants du risque de maintien en incapacité. Compte tenu de la composition des données du portefeuille, l'âge a finalement été retenu comme unique variable pour la construction de la loi d'expérience.

Construction de la loi d'expérience

Pour modéliser la loi d'expérience, l'utilisation des méthodes non paramétriques a été privilégiée. Ce choix est justifié par la non connaissance à priori de la loi théorique. L'estimateur de Kaplan-Meier a été ainsi retenu pour la construction des taux bruts de sortie d'incapacité. L'avantage de cette méthode est sa capacité à tenir compte de données incomplètes et plus particulièrement des censures à droite.

Les intervalles de confiance à 95% des taux de sortie ont été construits en se basant sur l'estimateur de Greenwood. La fiabilité du modèle a été validée au vu de la taille de l'intervalle de confiance sur les 12 premiers mois d'ancienneté. Par ailleurs, certaines irrégularités des taux de sortie ont été remarquées notamment pour les âges les moins élevés. Un lissage de ces taux par la méthode de Whittaker-Henderson a été effectué pour les diminuer.

Une fois l'ajustement validé statistiquement et graphiquement, une vérification de la bonne adéquation du modèle aux données du portefeuille a été réalisée. Pour cela, le nombre de sorties d'incapacité observées a été comparé aux valeurs prédites par le modèle. Le ratio de sorties modélisées par rapport aux sorties réelles était de 97,2%. Le principe de prudence a été ainsi respecté.

En raison de la volumétrie des données, les taux de sortie d'incapacité ont été modélisés jusqu'à 12 mois d'ancienneté et pour un âge d'entrée en incapacité limité à 60 ans. Pour étendre la loi d'expérience jusqu'à 36 mois d'ancienneté et au-delà de 60 ans, un raccordement a été effectué avec une table de référence. La table BCAC 2010 a été privilégiée pour réaliser ce prolongement. En effet, cette table est toujours en vigueur alors que la table BCAC 2013 n'a pas encore été homologuée depuis sa construction. De plus, la loi de maintien de 2010 est plus proche des observations du portefeuille étudié que celle de 2013 (la table de 2013 surestime la durée de maintien en incapacité par rapport à celle de 2010 pour les anciennetés faibles et sous-estime légèrement cette durée pour les anciennetés élevées).

Mise à l'épreuve de la loi d'expérience

Graphiquement, il a été constaté que la courbe de la loi d'expérience est toujours en dessous de celle de la loi réglementaire pour tous les âges d'entrée en incapacité. Ce résultat a confirmé que la table du BCAC surestime bien la durée de maintien en incapacité par rapport à la table d'expérience.

A partir de la loi d'expérience, les provisions mathématiques pour le versement d'une prestation de 1€ ont été calculées. Un barème de provisionnement pour un taux d'actualisation nul a été mis en place par la suite. Une comparaison avec le barème de provisionnement de la loi du BCAC a mis en évidence que les provisions calculées avec la loi d'expérience sont systématiquement moins élevées que celles calculées avec la loi réglementaire pendant les 12 premiers mois. Les courbes d'évolution des PM se superposent au-delà des 12 mois par construction.

Afin d'étudier l'impact de changement de table sur le provisionnement, une analyse boni/mali a été effectuée en comparant la variation des provisions mathématiques de maintien en incapacité au titre de la survenance 2019 entre les inventaires fin 2019 et fin 2020 aux prestations réellement payées en 2020. En conclusion, la loi d'expérience a diminué le montant des provisions mathématiques et a permis de mieux estimer la réalité du portefeuille.

Par ailleurs, les résultats de l'utilisation de la loi d'expérience dépendent fortement de la règle de sélection des encours. A ce titre, une adaptation de cette règle a été proposée au terme de ce mémoire. La solution proposée consiste à utiliser la loi d'expérience avec un décalage de la date d'observation des sinistres de 3 mois.

En effet, avec une sélection des encours au 31/03/N+1, la table du BCAC surestime la charge réelle de 61% (vs. 160% avec une sélection des encours au 31/12/N), alors que la loi d'expérience surestime la charge réelle de 8% seulement (vs. 83% avec la règle en vigueur de sélection des encours).

Cette proposition permet d'atteindre l'objectif attendu, à savoir fournir une estimation réaliste du risque tout en respectant les contraintes opérationnelles de production des comptes clients.

Executive summary

Context and issues

To honor their commitments to their policyholders, insurers must set up reserves representative of their risk exposure.

In foresight insurance, the constitution of technical reserves for claims in payment is a challenge for insurers. These reserves are calculated to deal with future payments of current benefits. They require probabilities calculation and updating of financial flows.

As part of the production of customer accounts, the reserving of the work stoppage risk is currently done using the BCAC's tables. These last are made on the portfolios of the main french insurance companies.

However, the use of these tables can lead to overestimate the technical reserves in certain conditions. This is the case of collective foresight insurance contracts which have a disability guarantee with a short deductible. Within this scope, recurring and notable surpluses are well-known for the risk of disability maintenance.

The portfolio studied in this brief concerns continuous deductible contracts. The period of which does not exceed 10 days. This is a special feature in several collective contracts insured by AG2R La Mondiale. This specificity is linked to the obligations of certain branch schemes.

When establishing the income statements, a large deficit C / P ratio is systematically noticed on the last occurrence presented. This indicator does not reflect reality. In consequence, it does not allow conclusions of possible technical imbalance of the disability risk and more generally of the insured plan.

In this perspective, the will is to build a disability maintenance table more suitable for the loss experience of this portfolio. The objective is to assess the corresponding reserves more precisely and therefore to ensure better management of the schemes concerned.

Data used

The portfolio is made up of collective contracts with direct management resulting from various National Collective Agreements. The initial data base contains 107,494 observations. Each observation represents a work stoppage that occurred between 10/15/1998 and 12/31/2018. The data is observed on 12/31/2019. Observing the data one year later from the last date of occurrence allow avoiding the problems of late declarations and having almost complete data. In addition, calculations and comparisons of technical reserves were carried out on claims from the 2019 occurrence to test the robustness of the model.

After an important step of data preparation, a statistical analysis of the portfolio was conducted. It showed that the studied population is mainly female. It is dominated by non-executives, married people, working in the health sector. The main cause of disability is the illness. The mean onset age of the disability is around 41 years old. Claims are closed in the first year 97% of the time.

The statistical results expose an heterogeneity of the length of the shutdown. To go further in the segmentation of risk profiles, the random forest method was applied. It highlighted that the salary, age, deductible, and type of work stoppage are discriminant parameters in the risk of disability

maintenance. Given the composition of the portfolio, age was finally retained as the sole segmentation variable for the construction of the experience law.

Construction of the experimental law

To model the experimental law, non-parametric methods were chosen. This is justified by the non-knowledge of the theoretical law. The Kaplan-Meier estimator was thus used for the construction of the crude discharge rates from disability. The advantage of this method is its ability to take into account incomplete data and especially right-wing censorship.

The 95% level of confidence intervals for exit rates were constructed with the Greenwood estimator. The reliability of the model was demonstrated given the size of the confidence interval over the first 12 months of seniority. In addition, some irregularities in the exit rates were noted, in particular for the youngest ages. These rates were smoothed using the Whittaker-Henderson method to reduce them.

The adjustment was validated both statistically and graphically. Then, a check of the model suitability to the portfolio data was made. To do this, the number of exits from disability observed were compared to the predicted ones. The number of modeled exits reported to actual exits was 97.2%. The principle of prudence was thus respected.

Due to the volumetry of the data, the exit disability rates were modeled up to 12 months of seniority. The entry age is limited to 60 years. To extend the law of experience up to 36 months of seniority and beyond 60 years, a connection was made with a reference table. The BCAC 2010 table was chosen to achieve this extension. Indeed, this table is still used while the BCAC 2013 table has not yet been approved since its construction. In addition, the 2010 maintenance law is closer to the observations of the portfolio studied than 2013 (the 2013 table overestimates the duration of disability maintenance compared to the 2010 table for low seniority and slightly underestimates this duration for senior staff).

Testing the law of experience

Graphically, it has been observed that the curve of the experience law is always below the regulatory law regardless of the entry's age into disability. This confirmed that the BCAC table overestimates the duration of disability compared to the experience table.

Based on the experimental law, the technical reserves for the payment of a benefit of € 1 were calculated. A reserving schedule for a zero update rate was subsequently constructed. A comparison with the reserving scale of the BCAC law showed that the technical reserves calculated with the experience law are systematically lower than those calculated with the regulatory law during the first 12 months. The reserves evolution curves are then superposed by construction beyond 12 months.

To study the impact of law changing on the reserving, a profit/loss analysis was done by comparing the variation of the technical reserves for maintenance in disability for the 2019 occurrence between the inventories at the end of 2019 and the end of 2020 with the services paid in 2020. In conclusion, the law of experience tends to reduce the technical reserves and became closer to the reality of the portfolio.

Furthermore, the results of using the experimental law strongly depend on the selection rule for outstandings. As such, an adaptation of this rule has been proposed at the end of this report. The proposed solution consists of using the experimental law with a 3-month delay in the observation date of claims.

Indeed, with a selection of outstanding at 03/31/ N + 1, the BCAC table overestimates the actual charge by 61% (vs. 160% with a selection of outstanding at 12/31/N). While the law of experience overestimates the actual charge by only 8% (vs. 83% with the current rule for selecting outstandings).

This proposal provides a realistic estimate of the risk while respecting the operational constraints of producing income statements. The target was thus achieved.