

Mémoire présenté pour l'obtention du DUAS et l'admission à l'Institut des Actuaires**Le 16 décembre 2022**Par : HAMZA KahinaTitre: Étude de l'hypothèse d'indépendance entre années de survenance dans le
risque de réserve d'un modèle interne partielConfidentialité : NON OUI Durée : 1 an 2 ans 3 ans 4 ans 5 ans*Membres du jury de l'IA :*S. JARRIER
D. DUBOIS*Entreprise : Groupama Assurances
Mutuelles**Directeur de mémoire (entreprise) :**Membres du jury de l'Unistra :*J. BERARD
E. BIRMELE
A. COUSIN
P.-O. GOFFARD
M. MAUMY-BERTRAND*Nom : MELIAND Charles*

Signature du responsable entreprise

Secrétariat : Mme Stéphanie Richard

Signature du candidat

REMERCIEMENTS

Avant tout développement sur cette expérience professionnelle, il apparaît opportun de commencer ce mémoire par des remerciements, à ceux qui m'ont beaucoup appris au cours de cette année d'alternance ainsi qu'à ceux qui m'ont soutenu durant la rédaction de mon mémoire.

Tout d'abord je tiens à remercier Charles MELIAND, responsable de l'équipe Modèle Interne Non-Vie, qui m'a encadré et accompagné en tant que maître d'apprentissage durant cette année d'alternance avec beaucoup de patience et de pédagogie. Je remercie l'ensemble des membres de l'équipe Modèle Interne Non-Vie pour leur accueil au sein de l'équipe, leur disponibilité ainsi que les échanges de qualité que j'ai pu avoir, leur soutien et leur implication tout au long de mon alternance mais également leur bonne humeur et la bonne ambiance présente au sein de l'équipe.

Je tiens également à remercier Sébastien KUNTZ, directeur de la Direction Actuariat Groupe, ainsi que l'ensemble des membres de la direction avec qui j'ai pu échanger sur divers thèmes et qui ont contribué au bon déroulement de mon année d'alternance.

Je tiens ensuite à remercier Thomas LALLEMENT, mon tuteur académique, pour sa disponibilité et son suivi durant la rédaction de mon mémoire, pour ses conseils et réponses lorsque j'en avais besoin.

Enfin, un grand remerciement à mes parents ainsi que mes frères et sœurs, qui m'ont soutenu et accompagné moralement et physiquement tout au long de ma formation d'actuaire.

RÉSUMÉ

Chez Groupama, le SCR de souscription non-vie se calcule à l'aide d'un modèle interne partiel qui fait l'objet d'une revue annuelle. Lors de la dernière revue, l'ACPR s'est intéressée au module de risque de réserves compris dans le module de risque de souscription non-vie. Le calcul du SCR de risque de réserves au sein du modèle interne de Groupama repose sur l'estimation de la volatilité à un an à l'aide de la méthode de Merz-Wuthrich. Cependant, cette méthode repose sur des hypothèses contraignantes comme l'hypothèse d'indépendance entre années de survenance. Cette hypothèse étant rejetée, une solution permettant de tenir compte de cette dépendance lors de l'estimation de la volatilité à un an a été demandée lors de la dernière revue.

Ce mémoire a pour objectif d'étudier la dépendance pouvant exister entre les années de survenance au sein du risque de réserves. Nous présenterons le modèle interne construit par Groupama dans une première partie en nous focalisant sur la méthode de calcul du SCR de risque de réserves. L'ensemble des concepts et méthodes théoriques utilisés pour mener notre étude seront présentés dans une seconde partie dans laquelle nous étudierons dans un premier temps, l'impact d'un facteur pouvant être à l'origine du rejet du test d'indépendance : l'inflation. Parmi les différentes approches, nous étudierons en premier lieu la corrélation présente au sein des triangles de données ainsi que l'impact d'ajout de corrélation par différentes méthodes. La seconde approche tourne autour du changement de modèle Chain-Ladder vers l'application d'un modèle linéaire généralisé (GLM) Tweedie. L'ensemble de ces méthodes seront appliquées et étudiées sur deux lignes d'activité : medical expenses (med.exp.) et income protection (inc.pr.) dont les résultats et analyses figureront dans une troisième et dernière partie.

Mots clés :

Modèle interne partiel, indépendance, corrélation, risque de réserves

ABSTRACT

At Groupama, the non-life underwriting risk SCR is calculated using a partial internal model that is reviewed annually. During the last review, the ACPR (Prudential Control and Resolution Authority) was interested in the reserve risk module included in the non-life underwriting risk module.

The calculation of the reserve risk SCR within Groupama's internal model is based on the estimation of the one-year volatility using the Merz-Wuthrich method. However, this method is based on restrictive assumptions such as the assumption of independence between years of occurrence. As this assumption was rejected, a solution allowing to consider this dependence when estimating the one-year volatility was requested during the last review.

The objective of this paper is to study the dependence that may exist between years of occurrence within the reserve risk. We will present the internal model built by Groupama in a first part by focusing on the method of calculation of the SCR of the reserve risk. All the theoretical concepts and methods used to conduct our study will be presented in a second part. Among these methods, we will first study the correlation present within the data loss triangles as well as the impact of adding correlation by different methods. The second approach revolves around the change from a Chain-Ladder model to the application of a Tweedie generalized linear model (GLM). Finally, we will study the impact of a factor that can be at the origin of the rejection of the independence test: the inflation. All these methods will be applied and studied on two lines of business: medical expenses (med.exp.) and income protection (inc.pr.) whose results and analysis will be presented in a third and final part.

Key words:

Partial internal model, independence, correlation, reserve risk

NOTE DE SYNTHÈSE

Contexte

Afin de mieux refléter sa structure et son profil de risque, Groupama a fait le choix de développer un Modèle Interne Partiel (MIP) pour le calcul du SCR de souscription non-vie qui comporte les risques de primes et de réserve. Son utilisation a été approuvée par l'Autorité de Contrôle Prudentielle et Réglementaire (ACPR) et l'assureur doit être capable de justifier toutes les hypothèses sous-jacentes auprès de cette dernière. Chaque année, une revue du modèle interne est effectuée suivant laquelle l'ACPR émet certaines recommandations.

Lors du dernier rapport de l'ACPR, il a été demandé d'établir une gouvernance en cas de rejet des tests d'hypothèses sous-jacentes à nos modèles et plus particulièrement sur l'hypothèse d'indépendance entre années de survenance dans le calcul du SCR du risque de réserves sous-jacente à l'utilisation de la méthode Chain Ladder.

Pour ce faire, nous allons étudier et analyser l'impact de la prise en compte de la dépendance entre années de survenance au sein des triangles de données. Notre étude suivra les différents points ci-dessous :

- Test d'hypothèses du modèle
- Etude d'un facteur calendaire
- Ajout de dépendance entre années de survenance
- Changement de méthode de provisionnement déterministe

Calcul de la volatilité autour du risque de réserve

Le SCR relatif au risque de réserves est obtenu selon la formule suivante :

$$SCR_{res} = BE_0 \cdot \left(\frac{\exp(q \cdot \sqrt{\ln(1 + \phi^2)})}{\sqrt{1 + \phi^2}} - 1 \right)$$

Où :

- BE_0 : provision pour sinistres best estimate en date $t = 0$
- ϕ : volatilité à 1 an relative au risque de réserves
- q : quantile à 99,5% d'une loi normale centrée réduite

Dans le cadre de notre étude, nous allons nous intéresser au calcul du paramètre ϕ correspondant à la volatilité à 1 an relative au risque de réserves. Il existe deux méthodes d'estimation de la volatilité à un an :

- Par la formule fermée de Merz-Wuthrich
- Par procédure Bootstrap

Groupama a choisi de retenir la méthode de réplication du modèle de Merz Wuthrich par la procédure Bootstrap. Ces deux méthodes visent à évaluer la volatilité du Claim Development Result (CDR) qui correspond à la différence entre la charge ultime vue en date $t = 0$ et la charge ultime vue en date $t = 1$, sur les survenances antérieures à la date $t = 0$. L'estimation des charges ultimes repose sur la méthode de provisionnement de Mack Chain-Ladder. Cette dernière repose sur trois hypothèses :

- (H1) : Les années de survenance i sont indépendantes entre elles
- (H2) : Pour une année de survenance i fixée, $\forall j, 1 \leq j \leq N - 1$, il existe un facteur de développement f_j tel que : $E(C_{i,j+1}|C_{i,j}) = f_j C_{i,j}$
- (H3) : La variance conditionnelle des charges (ou paiements) est telle que : $\forall j, 1 \leq j \leq N - 1$ et $\forall i, 1 \leq i \leq N$, $Var(C_{i,j+1}|C_{i,j}) = \sigma_j^2 C_{i,j}$

Par la suite nous nous focaliserons sur l'hypothèse d'indépendance entre années de survenance (H1).

Test d'hypothèse d'indépendance entre années de survenance

Le test d'indépendance entre années de survenance repose sur l'absence d'effets calendaire au sein du triangle des facteurs de développement individuels proposé par Thomas Mack. C'est une hypothèse forte qui se voit souvent rejetée. Etant le cas des LoB medical expenses et income protection, nous avons choisi de mener notre étude autour de ces dernières.

LoB	Z	E(Z)	Var(Z)	Intervalle de confiance		Hypothèse
med.exp	69	78,5	19,5	69,8	87,1	Rejetée
inc.pr	106	122,5	28,4	112,0	132,9	Rejetée

Résultats du test d'hypothèse d'indépendance entre années de survenance

Retraitement d'un facteur calendaire : l'inflation

Dans un premier temps, nous avons tenté de trouver la cause du rejet du test d'indépendance. Le test d'hypothèse étant basé sur l'absence d'effets calendaires, l'inflation peut être à l'origine de la présence de tels effets au sein de nos triangles et par conséquent responsable du rejet du test d'indépendance entre années de survenance.

Pour cela, nous avons retraité les effets de l'inflation des triangles de paiements en appliquant un facteur multiplicatif sur chaque diagonale de nos triangles. Ce facteur permet de rapporter les montants de chaque diagonale vue en année 2021 ce qui correspond à une « mise en as-if ».

Après avoir appliqué cette transformation, nous appliquons à nouveau le test d'hypothèse d'indépendance entre années de survenance qui, théoriquement, devrait être accepté.

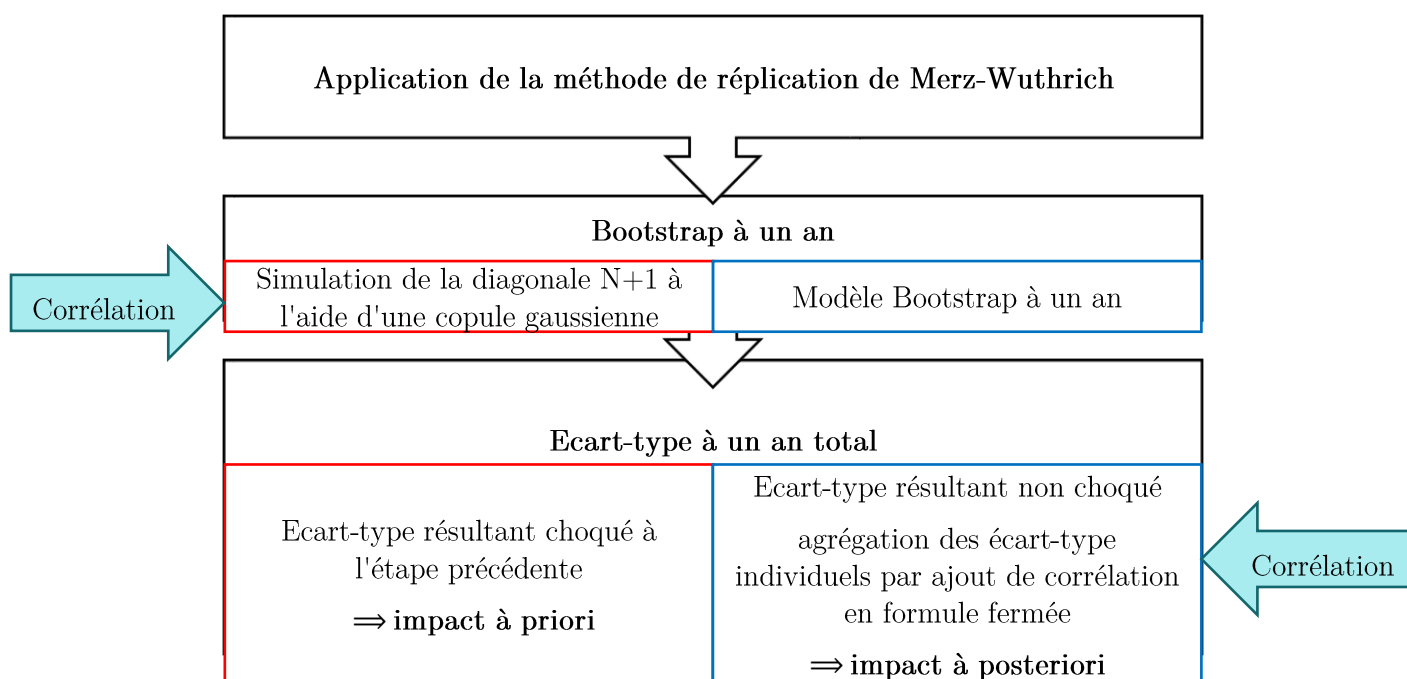
LoB	Z	E(Z)	Var(Z)	Intervalle de confiance		Hypothèse
med.exp	34	121,6	28,2	111,2	132,0	Rejetée
inc.pr	41	122,4	28,4	112,0	132,9	Rejetée

Résultats du test d'hypothèse d'indépendance entre années de survenance sur les triangles « as-if » 2021

Le retraitement de l'inflation dans le triangle des paiements ne supprime pas les effets calendaires à l'origine du rejet du test d'indépendance.

Ajout de dépendance entre années de survenance

Dans un second temps, nous avons appliqué la méthode actuelle utilisée par Groupama en ajoutant de la corrélation entre années de survenance suivant deux méthodes :



Les valeurs de corrélation testées lors des deux approches sont identiques (25%, 50%, 75%, 100%) et ont été choisies de manière arbitraire afin d'effectuer une première estimation.

Les volatilités à un an obtenues après la prise en compte de la corrélation entre années de survenance sont présentées ci-dessous :

Impact 25%	MSEP central	MSEP à posteriori	MSEP à priori
med.exp.	14,0	14,4	8,7
inc.pr.	21,5	26,8	31,5

Impact 50%	MSEP central	MSEP à posteriori	MSEP à priori
med.exp.	14,0	14,8	9,1
inc.pr.	21,5	31,7	39,3

Impact 75%	MSEP central	MSEP à posteriori	MSEP à priori
med.exp.	14,0	15,1	9,5
inc.pr.	21,5	36,0	45,9

Impact 100%	MSEP central	MSEP à posteriori	MSEP à priori
med.exp.	14,0	15,5	9,8
inc.pr.	21,5	39,8	51,7

Comparaison des MSEP à un an (en M€)

Nous remarquons que les résultats obtenus diffèrent en fonction de la LoB étudiée et de la méthode d'ajout de corrélation appliquée. La méthode à priori permet d'obtenir une volatilité plus élevée seulement sur la LoB income protection.

Changement de méthode de provisionnement déterministe

La dernière approche étudiée est le changement de méthode de provisionnement déterministe utilisée dans le modèle interne partiel de Groupama.

La méthode de Merz-Wuthrich repose sur l'utilisation de la méthode Chain-Ladder. Nous avons appliqué la méthode de provisionnement déterministe reposant sur l'utilisation d'un modèle linéaire généralisé (GLM).

Les provisions Best Estimate (BE) ainsi que les volatilités à un an obtenues après changement de méthode de provisionnement déterministe sont présentées ci-dessous :

med.exp.	Central	GLM Poisson	GLM Tweedie
Provisions BE (k€)	225 071	225 201	224 974
MSEP (k€)	14 018	5 855	34 307

inc.pr.	Central	GLM Poisson	GLM Tweedie
Provisions BE (k€)	442 355	442 396	443 308
MSEP (k€)	17 350	15 080	25 084

Provisions BE et écart-type à un an

Nous remarquons que pour les deux LoB, les résultats varient dans le même sens pour chaque GLM étudié. Le GLM Tweedie permet d'obtenir une volatilité à un an plus élevée que la méthode utilisée actuellement.

Conclusion

Notre étude consiste à étudier l'hypothèse d'indépendance entre années de survenance et trouver une alternative à son rejet.

Nous avons constaté que l'inflation n'est pas à l'origine de la présence d'effets calendaires au sein de nos triangles des paiements. La première approche consiste à appliquer le modèle actuel utilisé par Groupama à savoir, la méthode du Bootstrap à un an, sans tenir compte du rejet de l'hypothèse étudiée mais en la compensant par un ajout de corrélation. La seconde approche consiste à changer la méthode de provisionnement déterministe qui, dans le modèle actuel, correspond à la méthode Chain Ladder, par l'utilisation de modèle linéaire généralisé.

Enfin, les travaux de ce mémoire contribuent à l'amélioration continue du modèle interne partiel de Groupama. Des études supplémentaires et approfondies permettant de vérifier si ces différentes méthodes reflètent une réalité au sein de Groupama.

EXECUTIVE SUMMARY

Background

In order to better reflect its structure and risk profile, Groupama has chosen to develop a Partial Internal Model (PIM or MIP in French) for the calculation of the non-life underwriting SCR which includes premium and reserve risks. Its use has been approved by the “Autorité de Contrôle Prudentielle et Réglementaire (ACPR or Prudential and Regulatory Control Authority in English)” and the insurer must be able to justify all underlying assumptions to the ACPR. Each year, a review of the internal model is carried out, following which the ACPR issues certain recommendations.

During the last ACPR report, we were asked to establish governance in case of rejection of the hypothesis tests underlying our models and more particularly on the assumption of independence between years of occurrence in the calculation of the SCR of the reserve risk underlying the use of the Chain Ladder method.

To do so, we will study and analyze the impact of the inclusion of the dependence between years of occurrence within the data triangles. Our study will follow the different points below:

- Model hypothesis testing
- Study of a calendar factor
- Addition of dependency between years of occurrence
- Change of deterministic provisioning method

Calculation of the volatility around the reserve risk

The SCR for reserve risk is obtained using the following formula:

$$SCR_{res} = BE_0 \cdot \left(\frac{\exp(q \cdot \sqrt{\ln(1 + \phi^2)})}{\sqrt{1 + \phi^2}} - 1 \right)$$

With :

- BE_0 : best estimate claims provision at date $t=0$
- ϕ : 1-year volatility related to reserve risk
- q : 99.5% quantile of a standard normal distribution

In the context of our study, we will be interested in the calculation of the parameter ϕ corresponding to the 1-year volatility relative to the risk of reserves. There are two methods for estimating the 1-year volatility:

- With the Merz-Wuthrich closed formula
- With the Bootstrap Method

Groupama has chosen to use the replication method of the Merz Wuthrich model by the Bootstrap method. These two methods aim to evaluate the volatility of the Claim Development Result (CDR) which corresponds to the difference between the ultimate cost of claims seen at date $t=0$ and the ultimate cost of claims seen at date $t=1$, on events prior to date $t=0$. The estimation of ultimate cost of claims is based on the Mack Chain-Ladder provisioning method. This method is based on three assumptions :

- (H1): The years of occurrence i independent of each other
- (H2): For a fixed year of occurrence i , $\forall j, 1 \leq j \leq N - 1$, there exists a development factor f_j such as: $E(C_{i,j+1}|C_{i,j}) = f_j C_{i,j}$
- (H3): The conditional variance of payments is such as : $\forall j, 1 \leq j \leq N - 1$ et $\forall i, 1 \leq i \leq N$, $Var(C_{i,j+1}|C_{i,j}) = \sigma_j^2 C_{i,j}$

In the following, we will focus on the hypothesis of independence between years of occurrence (H1).

Test of the hypothesis of independence between years of occurrence

The test of independence between years of occurrence is based on the absence of calendar effects within the triangle of individual development factors proposed by Thomas Mack. This is a strong hypothesis that is often rejected. Being the case of LoB medical expenses and income protection, we have chosen to conduct our study around the latter.

LoB	Z	E(Z)	Var(Z)	Confidence range		Hypothesis
med.exp	69	78,5	19,5	69,8	87,1	Rejected
inc.pr	106	122,5	28,4	112	132,9	Rejected

Results of the hypothesis test for independence between years of occurrence

Adjustment for a calendar factor: inflation

First, we tried to find the cause of the rejection of the independence test. Since the hypothesis test is based on the absence of calendar effects, inflation may be the cause of the presence of such effects in our triangles and therefore responsible for the rejection of the independence test between years of occurrence.

To do this, we adjusted for the effects of inflation on the payment triangles by applying a multiplicative factor to each diagonal of our triangles. This factor makes it possible to report the amounts of each diagonal seen in year 2021, which corresponds to an "as-if" adjustment.

After applying this transformation, we again apply the hypothesis test of independence between years of occurrence which, theoretically, should be accepted.

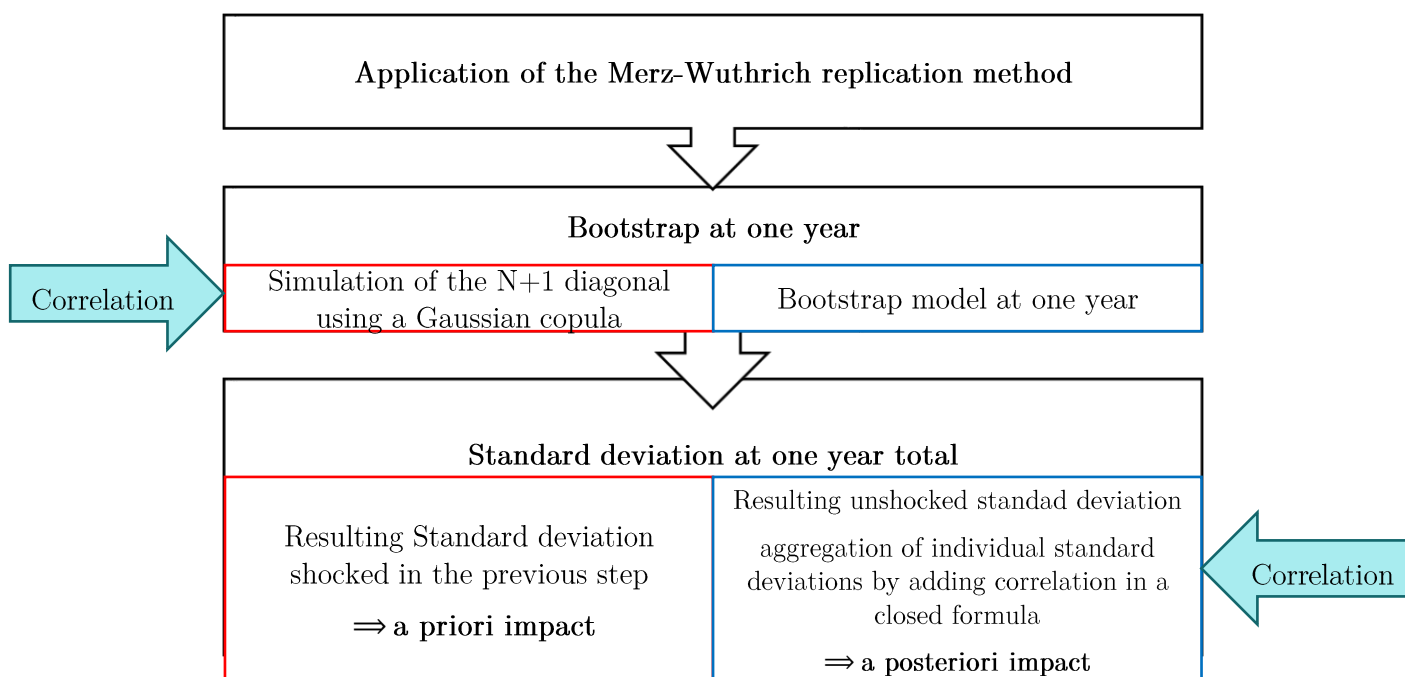
LoB	Z	E(Z)	Var(Z)	Confidence range		Hypothesis
med.exp	34	121,6	28,2	111,2	132,0	Rejected
inc.pr	41	122,4	28,4	112,0	132,9	Rejected

Results of the hypothesis test of independence between years of occurrence on the "as-if" triangles 2021

Adjusting for inflation in the payment's triangle does not remove the calendar effects that rejected the independence test.

Adding dependency between years of occurrence

In a second step, we applied the current method used by Groupama by adding correlation between years of occurrence following two methods:



The correlation values tested in both approaches are identical (25%, 50%, 75%, 100%) and were chosen arbitrarily to make an initial estimate.

The one-year volatilities obtained after taking into account the correlation between years of occurrence are presented below:

Impact 25%	Central MSEP	A Posteriori MSEP	A priori MSEP
med.exp.	14,0	14,4	8,7
inc.pr.	21,5	26,8	31,5

Impact 50%	Central MSEP	A Posteriori MSEP	A priori MSEP
med.exp.	14,0	14,8	9,1
inc.pr.	21,5	31,7	39,3

Impact 75%	Central MSEP	A Posteriori MSEP	A priori MSEP
med.exp.	14,0	15,1	9,5
inc.pr.	21,5	36,0	45,9

Impact 100%	Central MSEP	A Posteriori MSEP	A priori MSEP
med.exp.	14,0	15,5	9,8
inc.pr.	21,5	39,8	51,7

Comparison of one-year MSEP (in M€)

We note that the results obtained differ depending on the LoB studied and the correlation addition method applied. The a priori method allows to obtain a higher volatility only on the LoB income protection.

Change of deterministic provisioning method

The last approach studied is the change of the deterministic provisioning method used in Groupama's partial internal model.

The Merz-Wuthrich method is based on the use of the Chain-Ladder method. We applied the deterministic provisioning method based on the use of a generalized linear model (GLM).

Best Estimate (BE) provisions as well as the one-year volatilities obtained after the change of the deterministic provisioning method are presented below:

med.exp.	Central	Poisson GLM	Tweedie GLM
Provisions BE (k€)	225 071	225 201	224 974
MSEP (k€)	14 018	5 855	34 307

inc.pr.	Central	Poisson GLM	Tweedie GLM
Provisions BE (k€)	442 355	442 396	443 308
MSEP (k€)	17 350	15 080	25 084

BE provisions and standard deviation at one year

We note that for both LoBs, the results vary in the same direction for each GLM studied. The Tweedie GLM achieves higher one-year volatility than the method currently used.

Conclusion

Our study consists of studying the hypothesis of independence between years of occurrence and finding an alternative to its rejection.

We found that inflation is not the cause of the presence of calendar effects in our payment triangles. The first approach consists of applying the current model used by Groupama, namely the one-year bootstrap method, without considering the rejection of the hypothesis studied but compensating for it by adding correlation. The second approach consists in changing the deterministic provisioning method which, in the current model, corresponds to the Chain Ladder method, by using a generalized linear model.

Finally, the work in this thesis contributes to the continuous improvement of Groupama's partial internal model. Additional, in-depth studies enabling to verify whether these different methods reflect a reality within Groupama.

Table des matières

REMERCIEMENTS.....	2
RÉSUMÉ.....	3
ABSTRACT	4
NOTE DE SYNTHÈSE.....	5
EXECUTIVE SUMMARY	10
INTRODUCTION	17
Partie 1. CONTEXTE RÉGLEMENTAIRE ET MODÈLE INTERNE PARTIEL.....	18
A. Solvabilité 2.....	19
1. Définition	19
2. Les 3 piliers.....	19
3. Calcul du SCR.....	21
B. Modèle Interne Partiel.....	23
1. Calcul du SCR du risque de souscription non-vie.....	24
2. Calcul du SCR relatif au risque de réserves.....	28
3. Facteur de volatilité à un an.....	30
4. Tests des hypothèses de Merz-Wuthrich Chain-Ladder.....	38
Partie 2. THÉORIE DE LA DÉPENDANCE ET DES MÉTHODES ÉTUDIÉES.....	43
A. Prise en compte de l'inflation	44
1. Méthode de retraitement de l'inflation.....	44
2. Contexte économique actuel.....	45
B. Concepts mathématiques de dépendance.....	45
1. Fonction de répartition jointe.....	45
2. Notion et mesures de dépendance.....	46
3. Théorie des copules.....	49
C. Modèles linéaires généralisés (GLM).....	54
1. Définition	54
2. Provisionnement et GLM.....	56
3. Procédure Bootstrap	56
Partie 3. APPLICATION ET RÉSULTATS DES MÉTHODES ÉTUDIÉES.....	58
A. Données étudiées.....	59

B. Résultats des tests d'hypothèse	62
1. Hypothèse d'indépendance entre années de survenance (H1).....	62
2. Hypothèse de linéarité (H2)	62
3. Hypothèse de variance (H3).....	65
C. Triangles mis en « as-if »	68
D. Ajout de dépendance entre années de survenance	70
1. Choix de la copule.....	70
2. Méthodes d'ajout de dépendance	71
3. Corrélations entre années de survenance.....	78
4. Limites	81
E. Modèles linéaires généralisés	81
1. Résultats	81
2. Limites	95
CONCLUSION	96
ANNEXES.....	98
TABLE DES FIGURES	107
BIBLIOGRAPHIE.....	109

INTRODUCTION

Dans le cadre de la directive Solvabilité 2, les compagnies d'assurance doivent disposer d'un certain niveau de fonds propres correspondant au Solvency Capital Requirement (SCR) en plus des provisions Best Estimate (BE) au passif de leur bilan. Le SCR est le niveau des fonds propres requis que devra détenir la compagnie pour limiter sa probabilité de ruine à 0,5% sur 1 an. Autrement dit, il permet de couvrir les pertes liées à l'évènement de ruine qui voit sa réalisation une fois tous les 200 ans. Les provisions techniques Best Estimate sont les provisions permettant de couvrir le portefeuille actuel reflétant les engagements de l'assureur envers les assurés.

Pour calculer le SCR, la directive laisse le choix aux organismes d'assurance d'opter pour la méthode reflétant au mieux leur profil de risque. Ils peuvent opter pour les 3 méthodes suivantes :

- Formule standard
- Modèle interne (total ou partiel)
- USP

Groupama a fait le choix de développer un Modèle Interne Partiel (MIP) pour le calcul du SCR de souscription non-vie. Son utilisation a été approuvée par l'Autorité de Contrôle Prudentielle et Règlementaire (ACPR) et l'assureur doit être capable de justifier toutes les hypothèses sous-jacentes auprès de cette dernière. Chaque année, une revue du modèle interne est effectuée suivant laquelle l'ACPR émet certaines recommandations.

Lors du dernier rapport de l'ACPR, il a été demandé d'établir une gouvernance en cas de rejet des tests d'hypothèses sous-jacentes à nos modèles et plus particulièrement sur l'hypothèse d'indépendance entre années de survenance dans le calcul du SCR du risque de réserves sous-jacente à l'utilisation de la méthode Chain Ladder.

Ce mémoire a pour but d'analyser l'impact de la corrélation entre ces années au niveau de la volatilité à un an estimée pour le calcul du SCR du risque de réserves puis d'étudier et valider des méthodes alternatives afin de faire face au rejet de cette hypothèse.

Dans une première partie, nous présenterons la norme solvabilité 2 ainsi que le modèle interne partiel mis en place au sein de Groupama. Dans une deuxième partie nous définirons l'ensemble des modèles et concepts théoriques que nous utiliserons pour mener notre étude. Les analyses et résultats de l'application des méthodes présentées figureront dans une dernière partie.

**Partie 1. CONTEXTE
RÈGLEMENTAIRE ET MODÈLE
INTERNE PARTIEL**

A. Solvabilité 2

1. Définition

La solvabilité d'un organisme d'assurance est sa capacité à respecter les engagements pris auprès de ses assurés en détenant suffisamment de fonds propres pour faire face à la réalisation d'évènements imprévus.

Les compagnies d'assurance doivent disposer d'un montant de fonds propres supérieur aux provisions couvrant la totalité de leur portefeuille de risque.

La directive Solvabilité 2 est une réforme réglementaire européenne mise en place le 1er janvier 2016. Cette nouvelle directive a pour but d'améliorer l'évaluation et le contrôle des risques supportés par les organismes d'assurance afin que le niveau de fonds propres détenu soit suffisant et adéquat au profil de risques. Elle permet également de d'harmoniser la solvabilité sur le marché européen de l'assurance.

Cette directive se décompose en 3 piliers :

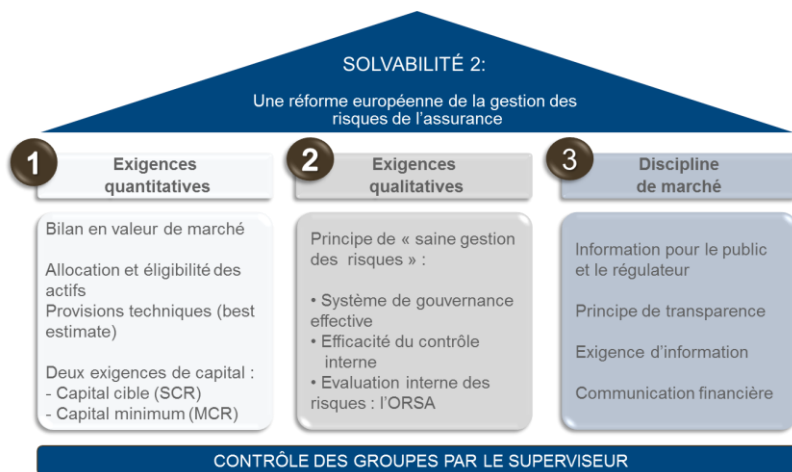


Figure 1 – Piliers de la directive Solvabilité 2

2. Les 3 piliers

a) Pilier 1 – Exigences quantitatives

Ce pilier permet d'estimer la richesse d'une compagnie et de mesurer ses risques. Toutes les règles de valorisation des actifs et des passifs y sont définies ainsi que les exigences en capital et la méthode retenue pour leur estimation.

Sous solvabilité 2, le bilan est établi en vision économique. Les actifs sont évalués en valeur de marché. Les provisions techniques sont évaluées de manière à avoir la meilleure estimation des flux futurs de

trésorerie et intègrent une marge de risque. C'est un bilan économique prudentiel qui sert à honorer ses engagements.

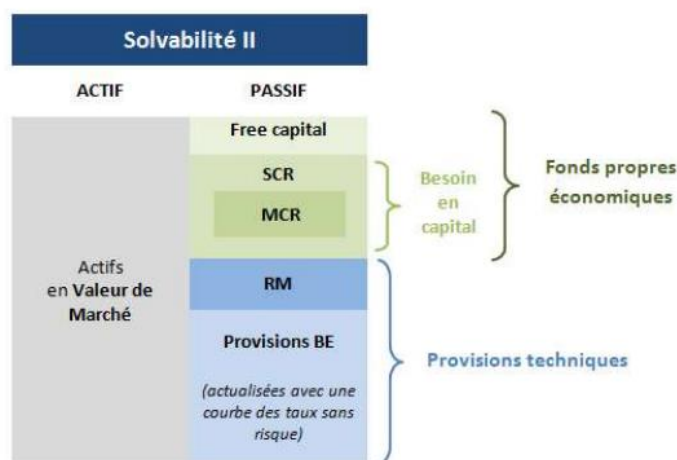


Figure 2 – Bilan en vision économique

Au niveau des fonds propres économiques, on retrouve :

- Le SCR (Solvency Capital Requirement) qui correspond au niveau de fonds propres permettant à une entreprise d'absorber des pertes imprévues importantes et d'offrir une sécurité raisonnable aux assurés et aux souscripteurs. Il permet de limiter sa probabilité de ruine à 0,5% sur un horizon d'un an.
- Le MCR (Minimum Capital Requirement) qui correspond au montant minimal de fonds propres à détenir pour pouvoir continuer à exercer l'activité d'assureur.

Au niveau des provisions techniques, on retrouve :

- Les Provisions Best Estimate (BE) correspondant à la valeur actuelle probable des flux de trésorerie futurs, déterminée sur la base d'informations actuarielles crédibles et d'hypothèses réalistes. Ces provisions reflètent la réalité du portefeuille assuré.
- Le Risk Margin (RM) couvre les risques liés à l'écoulement des passifs sur la totalité de leur durée. Il représente ainsi le montant qu'un repreneur éventuel de l'activité de la compagnie exigerait pour reprendre les engagements.

b) Pilier 2 – Exigences qualitatives

En complément des exigences quantitatives, ce pilier permet de connaître, maîtriser et piloter ses risques. L'évaluation interne des risques et de la solvabilité, plus couramment appelée l'ORSA (Own Risk and

Solvency Assessment) est un processus fournissant une évaluation à court et moyen terme des risques propres à la compagnie ainsi que du niveau de capital correspondant à les couvrir. Le système de gouvernance s'articule autour de quatre fonctions clés :

- La fonction de gestion des risques
- La fonction de conformité
- La fonction audit interne
- La fonction actuarielle

c) Pilier 3 – Discipline de marché

Ce dernier pilier vise à informer le public et les autorités de contrôle des exigences quantitatives et qualitatives des deux premiers piliers. Il a pour but d'accroître la transparence vis-à-vis du public. Les informations communiquées doivent refléter l'activité de la compagnie.

La communication des différentes informations est effectuée à l'aide des documents suivants :

- Le SFCR (Solvency and Financial Conditions Report) : rapport à destination du public synthétisant les principales évolutions de l'entreprise en termes de gouvernance, d'activité, de gestion du capital et de profil de risque.
- Le RSR (Regular Supervisory Reporting) : rapport à destination du superviseur visant à démontrer la solvabilité de l'entreprise).
- Les QRT (Quantitative Reporting Templates) : états quantitatifs destinés au superviseur et certains au public à une fréquence annuelle ou trimestrielle.

Dans la suite, nous allons nous intéresser au calcul du SCR établi et défini dans le premier pilier.

3. Calcul du SCR

Le calcul du SCR global se décompose en plusieurs modules de risques comme présentés ci-dessous. L'élément principal du SCR est le BSCR (Basic SCR). Il est atténué par le SCR de risque opérationnel qui permet de faire face au risque de perte résultant de procédures internes inadaptées ou défailtantes. L'ajustement permet de diminuer les excès de la formule standard dans le calcul du SCR de base.

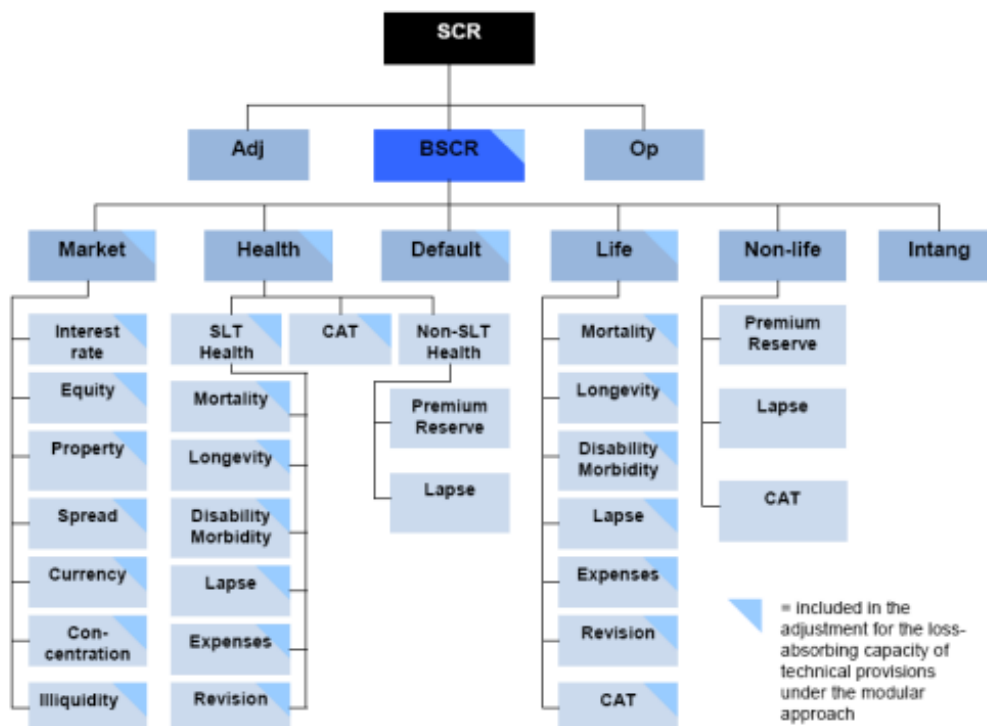


Figure 3 – Pieuvre des modules de la directive Solvabilité 2

Pour le calcul du SCR, les organismes d'assurance ont le choix entre 3 méthodes :

- Formule standard : cette méthode est dictée par la directive et la grande majorité des compagnies d'assurance l'utilise.
- Modèle interne : cette méthode de calcul est développée par les compagnies d'assurance afin de mieux refléter leur profil de risque. Elle doit cependant être étudiée et approuvée par l'Autorité de Contrôle Prudentielle et Réglementaire (ACPR).

Ce modèle peut être :

- Total : le modèle est entièrement développé par l'organisme d'assurance.
- Partiel : c'est une combinaison de la formule standard et du modèle interne total. Des paramètres du modèle peuvent être repris de la formule standard (matrice de corrélation par exemple)
- USP : cette méthode vise à modifier certains paramètres de calcul de la formule standard avec l'accord du superviseur.

Groupama a fait le choix de développer un modèle interne partiel afin de refléter au mieux son profil de risque.

B. Modèle Interne Partiel

Le modèle interne partiel de Groupama couvre le risque de souscription non-vie et le risque de souscription santé similaire à la non-vie (Non SLT). Le schéma ci-dessous distingue les modules de risques inclus dans le modèle interne partiel :

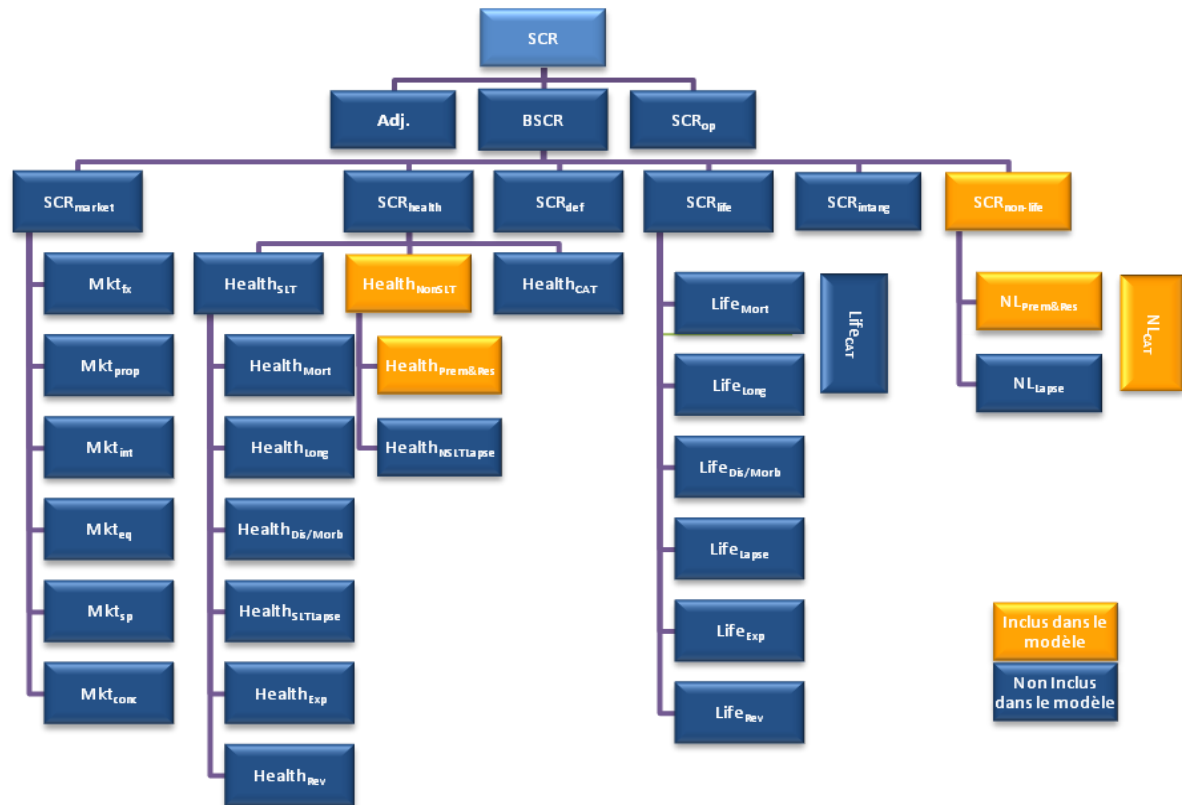


Figure 4 – Modules de risques couverts par le modèle interne de Groupama

La mesure de risque utilisée dans le MIP afin de calculer le SCR de souscription (non-vie et Non LST) est la Value-at-Risk (VaR) définie de la manière suivante :

$$VaR_{\alpha}(X) = \inf\{x | P(X \leq x) \geq \alpha\}$$

1. Calcul du SCR du risque de souscription non-vie

a) Définition du SCR de souscription non-vie

Le SCR de souscription est défini comme étant l'opposé de la Value-at-Risk (VaR) à 0,5% de la variation de la situation nette à un an. La situation nette est définie comme la différence entre les actifs et les passifs en valeur économique. La situation nette s'obtient à l'aide du calcul suivant :

$$SN = A - BE - RM$$

Où :

- A : actifs en valeur de marché
- BE : somme escomptée des flux futurs probables sur la base de la courbe des taux sans risque.
- RM : marge pour risque (Risk Margin)

Le SCR de souscription s'écrit de la manière suivante :

$$SCR = -VaR_{0.5\%}(SN_1 \cdot facteur_actu - SN_0)$$

Où :

- SN_1 : situation nette à un an (t=1)
- SN_0 : situation nette en date d'évaluation (t=0)
- $facteur_actu$: facteur d'actualisation

En supposant que l'organisme est soumis uniquement au risque de souscription non-vie, ce qui implique une variation d'actif nulle, et en incluant le facteur d'actualisation dans la notation de la situation nette, c'est-à-dire : $SN_1 \cdot facteur_actu = SN_1$, la variation de situation nette s'écrit :

$$SN_1 - SN_0 = -R - S - F + PE + [BE_0^{Pr} - BE_1^{Pr}] + RM_0 - RM_1$$

Où :

- R : boni/mali sur le stock de provisions représentant le risque de réserves
- S : charge sinistre relative aux nouveaux sinistres représentant le risque de prime
- F : ensemble des frais
- PE : prime émise pour la nouvelle année de souscription
- BE_i^{Pr} : provision pour primes en date $t = i$

La provision pour primes peut s'exprimer à l'aide du ratio combiné, des provisions pour primes non acquises et des primes futures. Nous pouvons écrire :

$$BE_i^{Pr} = PPNA_i \times \tau_i^{i+1} + PVFP_i \times (\tau_i^{i+1} - 1)$$

Où :

- $PPNA_i$: Provisions pour Primes Non Acquises comptabilisées en fin d'année i
- $PVFP_i$: Present Value of Future Premiums, assiette de primes futures à comptabiliser en fin d'année i
- τ_i^{i+1} : meilleure estimation, vue en fin d'année i, du ratio combiné net de réassurance de l'année i+1

La prime émise s'obtient en ajoutant à la prime acquise la variation de PPNA entre les dates $t = 0$ et $t = 1$. Elle peut donc s'exprimer de la manière suivante :

$$PE = PA + PPNA_1 - PPNA_0$$

Le résultat technique en vision économique vu en fin d'année 1 s'obtient grâce à la relation qui suit, à savoir :

$$TR = -(R + S + F) + PA$$

Avec l'ensemble des notations introduites précédemment, nous pouvons simplifier l'écriture de la variation de situation nette :

$$SN_1 - SN_0 = TR + (1 - \tau_1^2) \times (PNA_1 + PVFP_1) - (1 - \tau_0^1) \times (PNA_0 + PVFP_0) - \Delta RM$$

Où :

- $\Delta RM = RM_1 - RM_0$

Sous les hypothèses suivantes :

- La variation de marge pour risque est négligeable. ($\Delta RM = 0$)
- Les volumes sous risques de provision pour primes nettes de réassurance en fin d'année 1 ($PPNA_1$ et $PVFP_1$) sont déterministes.

Après simplification, la formule du SCR de souscription non-vie devient :

$$SCR = -VaR_{0.5\%} \left(TR + (\tau_0^1 - \tau_1^2) \times (PPNA_1 + PVFP_1) \right) + (1 - \tau_0^1) * (PPNA_0 + PVFP_0 - PPNA_1 - PVFP_1)$$

b) *Décomposition du SCR de souscription non-vie par ligne d'activité*

En assurance non-vie, la classification des risques est effectuée par « Line of Business (LoB) », définies dans les spécifications techniques fournies par l'EIOPA. Ces lignes d'activité regroupent les différents risques existants de sorte à construire des groupes de risques homogènes.

Il est possible d'exprimer la variation de situation nette en fonction des LoBs Solvabilité 2, ce qui nous amène à l'expression de SCR suivante :

$$SCR = -VaR_{0.5\%} \left(\sum_{i \in LoB} [TR_i + (\tau_{0,i}^1 - \tau_{1,i}^2) \times (PPNA_{1,i} + PVFP_{1,i})] \right) + \sum_{i \in LoB} [(1 - \tau_{0,i}^1) \times (PPNA_{0,i} + PVFP_{0,i} - PPNA_{1,i} - PVFP_{1,i})]$$

Où :

- $TR_i = PA_i - S_i - R_i - F_i$: résultat technique de la LoB i

Le SCR de souscription non-vie est composé :

- D'une partie aléatoire : $\sum_{i \in LoB} [TR_i + (\tau_{0,i}^1 - \tau_{1,i}^2) \times (PPNA_{1,i} + PVFP_{1,i})]$
- D'une partie déterministe : $\sum_{i \in LoB} [(1 - \tau_{0,i}^1) \times (PPNA_{0,i} + PVFP_{0,i} - PPNA_{1,i} - PVFP_{1,i})]$ que l'on va noter K par la suite.

Dans le modèle interne, la distribution complète de la variable aléatoire $\sum_{i \in LoB} [TR_i + (\tau_{0,i}^1 - \tau_{1,i}^2) \times (PPNA_{1,i} + PVFP_{1,i})]$ ne peut pas être obtenue par tirages de Monte Carlo. Une approche modulaire a été adoptée en distinguant deux risques :

- Le risque de réserves relatif à l'incertitude sur les montants et dates des paiements des sinistres déjà survenus.
- Le risque de prime relatif aux sinistres non encore survenus liés à la souscription pour l'année à venir.

Le SCR d'une LoB i :

$$SCR_i = VaR_{99.5\%} (R_i + S_i - PA_i + F_i - (\tau_{0,i}^1 - \tau_{1,i}^2) \times (PPNA_{1,i} + PVFP_{1,i})) + K$$

Les primes, les frais et le ratio combiné $\tau_{0,i}^1$ étant déterministes, nous pouvons les inclure dans la constante K pour finalement obtenir :

$$SCR_i = VaR_{99,5\%}(A) + K$$

Où :

- $A = R_i + S_i + \tau_{1,i}^2 \times (PPNA_{1,i} + PVFP_{1,i})$

En centrant les variables S_i et $\tau_{1,i}^2$ et en notant :

- σ_i^R : écart type de la variable R_i
- σ_i^P : écart type de la variable $S_i + \tau_{1,i}^2 \times (PPNA_{1,i} + PVFP_{1,i})$
- $\rho_{R,P}$: le coefficient de corrélation entre les variables R_i et $S_i + \tau_{1,i}^2 \times (PPNA_{1,i} + PVFP_{1,i})$

L'écart type de la variable A s'exprime comme suit :

$$\sqrt{(\sigma_i^R)^2 + (\sigma_i^P)^2 + 2 \cdot \rho_{R,P} \cdot \sigma_i^R \cdot \sigma_i^P}$$

Dans un cadre elliptique :

$$VaR_{99,5\%}(A - E(A)) = q \cdot \sqrt{(\sigma_i^R)^2 + (\sigma_i^P)^2 + 2 \cdot \rho_{R,P} \cdot \sigma_i^R \cdot \sigma_i^P}$$

Où :

- q : quantile à 99,5% d'une loi normale centrée réduite.

Dans le cas où la LoB i est soumise uniquement aux risques de prime et de réserve :

$$SCR_i = q \cdot \sqrt{(\sigma_i^R)^2 + (\sigma_i^P)^2 + 2 \cdot \rho_{R,P} \cdot \sigma_i^R \cdot \sigma_i^P} + K$$

Où :

- $K = -E(TR_i) + (1 - \tau_{0,i}^1) \times (PPNA_{0,i} + PVFP_{0,i} - PPNA_{1,i} - PVFP_{1,i})$

Le SCR de souscription non-vie a finalement trois composantes :

- $SCR_{res}^i = VaR_{99,5\%}(R_i) = q \cdot \sigma_i^R$ correspondant au SCR du risque de réserves.
- La variation de la sinistralité à 99,5% au-delà de la moyenne : $\widetilde{SCR}_{pr}^i = VaR_{99,5\%}([S_i + \tau_{1,i}^2 \times (PNA_{1,i} + PVFP_{1,i})] - [E(S_i) + E(\tau_1^2) \times (PNA_{1,i} + PVFP_{1,i})])$ qui diffère du SCR du risque de prime.
- La partie déterministe K .

Nous pouvons finalement écrire :

$$SCR_i = \sqrt{SCR_{res}^i{}^2 + \widetilde{SCR}_{pr}^i{}^2 + 2 \cdot \rho_{R,P} \cdot SCR_{res}^i \cdot \widetilde{SCR}_{pr}^i + K}$$

2. Calcul du SCR relatif au risque de réserves

Le risque de réserves correspond à l'incertitude relative à l'estimation des provisions pour sinistres à payer. Il représente ainsi le risque que les provisions constituées pour couvrir ces sinistres se révèlent insuffisantes au regard de leur coût réel. La directive Solvabilité 2 conduit à définir le risque de réserves comme le risque d'une déviation de la charge ultime des sinistres entre deux arrêts annuels successifs.

Le SCR relatif au risque de réserves à la date $t = 0$ représente le montant de fonds propres que l'organisme doit immobiliser pour couvrir le risque de réserves avec un niveau de confiance de 99,5% sur un horizon d'un an. Il peut ainsi s'écrire de la façon suivante :

$$SCR_{res} = SN_0 - VaR_{0,5\%}(SN_1)$$

Où :

- SN_i : situation nette en date $t = i$

En $t = 0$, la situation nette peut s'écrire :

$$SN_0 = A_0 - BE_0$$

Où :

- A_0 : valeur économique des actifs
- BE_0 : best estimate de sinistres

En $t = 1$, sous l'hypothèse que la compagnie est soumise uniquement au risque de réserves, la situation nette peut s'écrire :

$$SN_1 = A_1 - BE_1 = A_0 - P_1 - BE_1$$

Où :

- A_0 : valeur économique des actifs
- P_1 : paiement effectué durant l'année 1
- BE_1 : best estimate de sinistres

Le risque de taux n'étant pas inclus dans le modèle interne, les facteurs d'escomptes à horizon un an sont déterministes. Le SCR relatif au risque de réserves s'exprime donc comme suit :

$$SCR_{res} = A_0 - BE_0 - VaR_{0,5\%}(A_0 - P_1 - BE_1)$$

Les actifs A_0 étant déterministes, l'expression précédente devient :

$$SCR_{res} = -VaR_{0,5\%}(-(P_1 + BE_1)) - BE_0$$

Le Best Estimate BE_0 étant déterministe, nous pouvons modéliser le risque de réserves à l'aide la variable aléatoire suivante, représentant le taux de boni-mali entre les dates $t = 0$ et $t = 1$:

$$X_{res} = \frac{BE_1 + P_1}{BE_0}$$

Le SCR relatif au risque de réserves :

$$SCR_{res} = BE_0(VaR_{99,5\%}(X_{res}) - 1)$$

La variable aléatoire $BE_1 + P_1$ représente le boni-mali réalisé entre $t = 0$ et $t = 1$. En moyenne, les bonis-malis annuels sont nuls selon le caractère « best estimate » de la provision pour sinistres à payer ce qui implique que la variable $BE_1 + P_1$ est centrée sur BE_0 et par conséquent la variable aléatoire X_{res} est centrée sur 1.

La variable aléatoire $BE_1 + P_1$ supposée suivre une loi log-normale, cela implique que la variable X_{res} suit également une loi log-normale centrée sur 1 et d'écart type ϕ correspondant au coefficient de variation de $BE_1 + P_1$.

Sous ces hypothèses et notations, le SCR relatif au risque de réserves s'écrit :

$$SCR_{res} = BE_0 \cdot \left(\frac{\exp(q \cdot \sqrt{\ln(1 + \phi^2)})}{\sqrt{1 + \phi^2}} - 1 \right)$$

Où :

- BE_0 : provision pour sinistres best estimate en date $t = 0$
- ϕ : volatilité à 1 an relative au risque de réserves
- q : quantile à 99,5% d'une loi normale centrée réduite

Le calcul du SCR relatif au risque de réserves est effectué pour chaque LoB Solvabilité 2, présentées ci-dessous :

Non life UW risk	Health non SLT
1 : RC Automobile	1 : Frais de soins
2 : Dommages Automobile	2 : Income protection
3 : Marine, Aviation, Transport	
4 : Dommages aux Biens 4.1 : Dommage Ouvrage 4.2 : Hors Dommage Ouvrage	
5 : RC Générale 5.1 : RC Décennale 5.2 : Hors Décennale	
6 : Crédit	
7 : Protection juridique	
8 : Assistance	
9 : Divers	

Figure 5 – Lignes d’activité prises en compte au sein du modèle interne

Le SCR de réserve repose donc sur l’estimation de la volatilité à un an ϕ ainsi que les provisions pour sinistres best estimate BE_0 .

Dans la suite de l’étude, nous allons nous focaliser sur l’estimation de la volatilité à un an.

3. Facteur de volatilité à un an

a) Définition du facteur de volatilité à un an ϕ

La volatilité à un an de la variable X_{res} peut être définie comme le coefficient de variation de la variable $BE_1 + P_1$. Le coefficient de variation d’une variable est défini comme le rapport entre l’écart type de cette variable et sa moyenne.

La variable $BE_1 + P_1$ étant centrée sur BE_0 , nous pouvons exprimer ϕ comme suit :

$$\phi = \frac{\sigma(BE_1 + P_1)}{BE_0}$$

b) Estimation de la volatilité à un an

Afin d’estimer la volatilité à un an, il est possible de s’appuyer sur deux méthodes qui convergent vers le même résultat :

- Par la formule fermée de Merz-Wuthrich
- Par procédure Bootstrap

Ces deux méthodes visent à évaluer la volatilité du Claim Development Result (CDR) qui correspond à la différence entre la charge ultime vue en date $t = 0$ et la charge ultime vue en date $t = 1$, sur les survenances antérieures à la date $t = 0$. L'estimation des charges ultimes repose sur la méthode de provisionnement de Mack Chain-Ladder.

(1) Méthode Chain-Ladder

La méthode Chain-Ladder est une méthode déterministe de provisionnement. Elle permet de calculer le montant des provisions nécessaire à couvrir le portefeuille assuré. C'est la méthode la plus classique en provisionnement non-vie par sa simplicité de mise en œuvre. Elle s'applique aux triangles cumulés de charges comme de paiements.

Considérons un triangle de charges (ou paiements) cumulés avec N années de survenance et de développement de la forme suivante :

	1	...	j	...	N
1	$C_{1,1}$...	$C_{1,j}$...	$C_{1,N}$
⋮	⋮			⋮	
i	$C_{i,1}$		$C_{i,j}$		
⋮	⋮	⋮			
N	$C_{N,1}$				

Nous utiliserons les notations suivantes :

- i : année de survenance, $1 \leq i \leq N$
- j : année de développement, $1 \leq j \leq N$
- $C_{i,j}$: montant de la charge (ou paiement) cumulé de l'année de survenance i et l'année de développement j

La méthode Chain-Ladder repose sur les deux hypothèses suivantes :

- (H1) : Les charges cumulées $C_{i,j}$ des années de survenance i sont indépendantes. Autrement dit, les années de survenance sont indépendantes entre elles, $\forall i \neq i'$, $(C_{i,1}, \dots, C_{i,N})$ est indépendant de $(C_{i',1}, \dots, C_{i',N})$.
- (H2) : Pour une année de survenance i fixée, il existe une relation linéaire entre les charges cumulées. Autrement dit, $\forall j$, $1 \leq j \leq N - 1$, il existe un facteur de développement f_j tel que $\forall i$, $1 \leq i \leq N$:

$$C_{i,j+1} = f_j * C_{i,j}$$

Les facteurs de développement individuel $f_{i,j}$ sont calculés de la manière suivante : $\forall j, 1 \leq j \leq N - 1$ et $\forall i, 1 \leq i \leq N - 1$:

$$f_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}$$

Ces facteurs individuels constituent un triangle nommé le d-triangle.

L'estimation des facteurs de développement f_j est donnée par la relation qui suit :

$$\hat{f}_j = \frac{\sum_{i=1}^{N-j} C_{i,j+1}}{\sum_{i=1}^{N-j} C_{i,j}}, \quad j = 1, \dots, N - 1$$

A l'aide de ces facteurs et de l'hypothèse de linéarité (H2), nous pouvons compléter la partie inférieure du triangle des charges à l'aide de la relation suivante :

$$\hat{C}_{i,j+1} = \hat{f}_j * C_{i,j}, \forall i + j \geq N + 1$$

Les provisions individuelles pour une année de survenance i s'écrit :

$$\hat{R}_i = \hat{C}_{i,N} - C_{i,N-i+1}$$

La provision totale est donc la somme des provisions individuelles :

$$\hat{R} = \sum_{i=1}^N \hat{R}_i$$

(2) Méthode de Merz-Wuthrich en formule fermée

La méthode de Merz-Wuthrich est une méthode stochastique de provisionnement. Elle se base sur la méthode Chain-Ladder et permet de mesurer l'incertitude liée à l'estimation des provisions et à la charge en vision un an.

Cette méthode repose sur les mêmes hypothèses que la Chain-Ladder avec une hypothèse de variance en supplément. En reprenant les notations de la partie précédente, les hypothèses de la méthode de Merz-Wuthrich sont les suivantes :

- (H1) : Les années de survenance i sont indépendantes entre elles
- (H2) : Pour une année de survenance i fixée, $\forall j, 1 \leq j \leq N - 1$, il existe un facteur de développement f_j tel que : $E(C_{i,j+1}|C_{i,j}) = f_j C_{i,j}$
- (H3) : La variance conditionnelle des charges (ou paiements) est telle que : $\forall j, 1 \leq j \leq N - 1$ et $\forall i, 1 \leq i \leq N$, $Var(C_{i,j+1}|C_{i,j}) = \sigma_j^2 C_{i,j}$

Les paramètres σ_j , appelés sigmas de Mack, permettent de mesurer l'écart entre les facteurs de développement individuel $f_{i,j}$ et les facteurs de développement f_j . Mack propose une estimation sans biais de ces variances :

$$\begin{cases} \hat{\sigma}_j^2 = \frac{1}{N-j-1} \sum_{i=1}^{N-j} C_{i,j} (f_{i,j} - \hat{f}_j)^2, & 1 \leq j \leq N-2 \\ \hat{\sigma}_{N-1}^2 = \min\left(\frac{\hat{\sigma}_{N-2}^4}{\hat{\sigma}_{N-3}^2}; \min(\hat{\sigma}_{N-3}^2; \hat{\sigma}_{N-2}^2)\right) \end{cases}$$

La méthode de Merz-Wuthrich est une méthode non simulatoire permettant de mesurer la volatilité à un an par formule fermée. C'est une approche similaire à la méthode de Mack mais, contrairement à cette dernière, elle est adaptée au calcul du risque sur un horizon d'un an. Elle est également basée sur les CDR qui, pour rappel, sont les bonis-malis entre les dates $t = 0$ et $t = 1$ pouvant s'exprimer comme :

$$\widehat{CDR}_i(N+1) = \hat{C}_{i,N}^N - \hat{C}_{i,N}^{N+1}$$

Ou encore :

$$\widehat{CDR}_i(N+1) = \hat{R}_i^N - (\hat{Y}_{i,N-i} + \hat{R}_i^{N+1})$$

Où :

- $\hat{C}_{i,N}^N$: estimation de la charge ultime de l'année de survenance i vue en N
- $\hat{C}_{i,N}^{N+1}$: estimation de la charge ultime de l'année de survenance i vue en N+1
- \hat{R}_i^N : estimation des réserves pour l'année de survenance i vues en N
- \hat{R}_i^{N+1} : estimation des réserves pour l'année de survenance i vues en N+1
- $\hat{Y}_{i,N-i+1}$: estimation des paiements effectués entre N et N+1

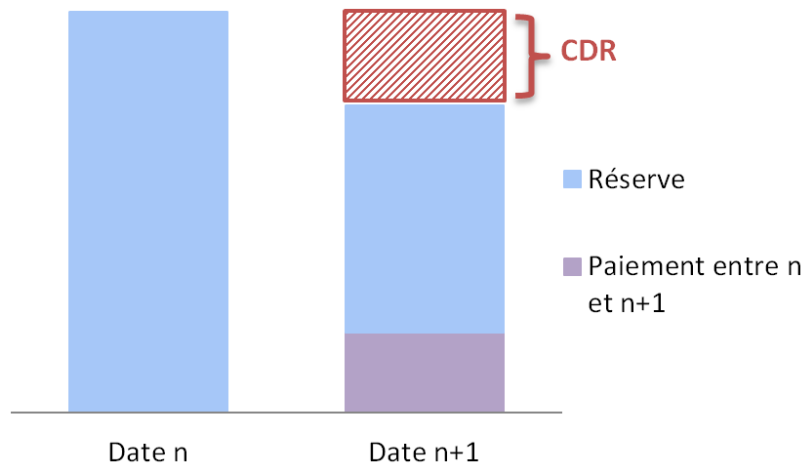


Figure 6 – Calcul du Claim Development Result

Il existe deux approches de calcul de volatilité à un an des CDR :

- Méthode prospective : l'erreur de prédiction du CDR mesure l'incertitude entre la valeur nulle et le CDR observable à la fin de l'année comptable N
- Méthode rétrospective : l'erreur de prédiction du CDR analyse la distance entre le CDR réel et le CDR observable.

L'approche retenue par le modèle interne partiel est l'approche prospective. Par conséquent, nous présenterons uniquement cette approche.

Afin de simplifier la lecture des formules, nous posons les éléments suivants pour une années de survenance i et une année de développement j :

- $S_j^N = \sum_{i=1}^{N-j-1} C_{i,j}$, $S_j^{N+1} = \sum_{i=1}^{N-j} C_{i,j}$
- $\hat{\Delta}_{i,N}^N = \frac{\hat{\sigma}_{N-i}^2 / \hat{f}_{N-i}^2}{S_{N-i}^N} + \sum_{j=N-i-1}^{N-1} \left(\frac{C_{N-j,j}}{S_j^{N+1}} \right)^2 \times \frac{\hat{\sigma}_j^2 / \hat{f}_j^2}{S_j^N}$
- $\hat{\Phi}_{i,N}^N = \sum_{j=N-i-1}^{N-1} \left(\frac{C_{N-j,j}}{S_j^{N+1}} \right)^2 \times \frac{\hat{\sigma}_j^2 / \hat{f}_j^2}{C_{N-j,j}}$
- $\hat{\Psi}_i^N = \frac{\hat{\sigma}_{N-i}^2 / \hat{f}_{N-i}^2}{C_{i,N-i}}$
- $\hat{\Gamma}_{i,N}^N = \hat{\Phi}_{i,N}^N + \hat{\Psi}_i^N$
- $\hat{\Lambda}_{i,N}^N = \frac{C_{i,N-i}}{S_{N-i}^{N+1}} \times \frac{\hat{\sigma}_{N-i}^2 / \hat{f}_{N-i}^2}{S_{N-i}^N} + \sum_{j=N-i-1}^{N-1} \left(\frac{C_{N-j,j}}{S_j^{N+1}} \right)^2 \times \frac{\hat{\sigma}_j^2 / \hat{f}_j^2}{S_j^N}$
- $\hat{Y}_{i,N}^N = \hat{\Phi}_{i,N}^N + \frac{\hat{\sigma}_{N-i}^2 / \hat{f}_{N-i}^2}{S_{N-i}^N}$

En vision prospective, la volatilité à un an par année de survénance i s'évalue de la manière suivante :

$$MSEP_{\widehat{CDR}_i(N+1)|D}(0) = E \left[(\widehat{CDR}_i(N+1) - 0)^2 | D \right]$$

Avec :

- $\widehat{CDR}_i(N+1)$: boni-mali observable entre les années comptables N et $N+1$
- D : ensemble d'informations connues au moment du calcul i.e. ensemble des éléments du triangle supérieur : $D = \{C_{i,j} \mid i+j \leq N+1\}$

En développant les calculs, la volatilité à un an s'exprime comme suit :

$$MSEP_{\widehat{CDR}_i(N+1)|D}(0) = (\hat{C}_{i,N}^N)^2 \times (\hat{\Gamma}_{i,N}^N + \hat{\Delta}_{i,N}^N)$$

L'estimateur de la variance du CDR s'obtient de la manière suivante :

$$\widehat{Var}(\widehat{CDR}_i(N+1)|D) = (\hat{C}_{i,N}^N)^2 \times \hat{\Psi}_i^N$$

La MSEP totale ne s'obtient pas par somme des MSEP par années de survénance mais par agrégation des MSEP individuelles suivant la formule ci-dessous :

$$MSEP(\hat{R}) = \sum_{i=1}^N MSEP_{\widehat{CDR}_i(N+1)|D}(0) + 2 \times \sum_{k>i>1} \hat{C}_{i,N}^N \times \hat{C}_{k,N}^N \times (\hat{\Lambda}_{i,N}^N + \hat{\Upsilon}_{i,N}^N)$$

La méthode de Merz-Wuthrich est décrite dans l'article « Modelling the claims development result for solvency purposes ».

Au sein du modèle interne, l'incertitude liée à l'estimation des provisions et de la charge ultime se fait sur un horizon d'un an à l'aide de la réplcation de la méthode de Merz-Wuthrich par Bootstrap.

(3) Méthode de Merz-Wuthrich par Bootstrap

La procédure Bootstrap permet d'obtenir la distribution empirique des provisions Best Estimate sur un horizon d'un an ainsi que des CDR. De plus, elle permet de prendre en compte certains retraitements spécifiques ne pouvant pas être pris en compte en formule fermée.

A partir des méthodes Chain-Ladder et Mack décrites dans les parties précédentes et en gardant les mêmes notations, nous pouvons calculer les résidus de Pearson correspondants aux résidus du modèle :

$$r_{i,j} = \frac{\sqrt{C_{i,j}} (f_{i,j} - \hat{f}_j)}{\hat{\sigma}_j}$$

Ces estimateurs étant biaisés, ils sont ajustés par un coefficient pour les rendre non biaisés :

$$r_{i,j} = \sqrt{\frac{N-j}{N-j-1}} \frac{\sqrt{C_{i,j}} (f_{i,j} - \hat{f}_j)}{\hat{\sigma}_j}$$

Ces résidus ajustés du biais sont centrés colonne par colonne respectivement par les facteurs de développement \hat{f}_j .

Le modèle Bootstrap suppose que les résidus sont indépendants et identiquement distribués (i.i.d).

Cette approche consiste à répéter B fois les étapes suivantes :

1. Rééchantillonnage des résidus de Pearson ajustés et centrés r_{ij}^* et construction du d-triangle des pseudo-facteurs de développement individuels à partir des paramètres issus de la méthode Mack Chain-Ladder et des résidus rééchantillonnés r_{ij}^* :

$$f_{i,j}^* = r_{ij}^* \frac{\hat{\sigma}_j}{\sqrt{C_{i,j}}} + \hat{f}_j$$

2. Calcul des pseudo-facteurs de développement à l'aide de la méthode Chain-Ladder pour obtenir une nouvelle cadence de paiement :

$$\hat{f}_j^* = \frac{\sum_{i=1}^{N-j} C_{i,j} \times f_{i,j}^*}{\sum_{i=1}^{N-j} C_{i,j}}$$

3. Simulation de la diagonale de l'année suivante (diagonale N+1) à l'aide d'une loi log-normale dont les paramètres sont déterminés à partir des montant cumulés de la diagonale précédente (diagonale N) et des pseudo-facteurs de développement issus de l'étape 2, $\forall 2 \leq i \leq N$:

$$\hat{C}_{i,N-i+2}^* \rightarrow \text{Lognormale}(\text{Mean} = \hat{f}_{N-i+1}^* C_{i,N-i+1} \mid \text{Variance} = \hat{\sigma}_{N-i+1}^2 C_{i,N-i+1})$$

4. Réévaluation des pseudo-facteurs de développement de Chain-Ladder, incluant la nouvelle diagonale N+1 :

$$\hat{f}_j^{*N+1} = \frac{\sum_{i=1}^{N-j} f_{i,j} C_{i,j} + \hat{C}_{N-j+1,j+1}^*}{\sum_{i=1}^{N-j} C_{i,j}}$$

5. Estimation des nouvelles charges ultimes avec la méthode de Chain-Ladder suivant la dernière diagonale simulée et des pseudo-facteurs de développement calculés à l'étape 4.

La variance des B estimations de charges ultimes permet d'évaluer la MSEP, ce qui nous donne, en considérant la racine de la MSEP, le numérateur de l'expression du coefficient de variation ϕ . Au sein du modèle interne de Groupama, le calcul de la volatilité à un an est réalisé suite à 20000 simulations de charges ultimes.

Le schéma ci-dessous reprend l'ensemble des étapes de la procédure Bootstrap.

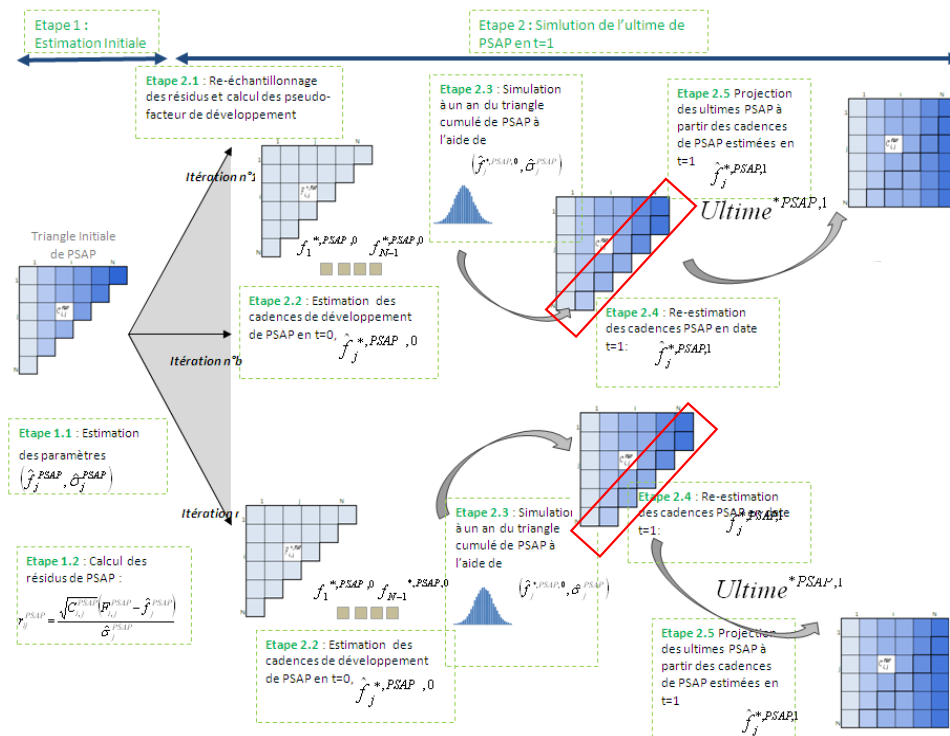


Figure 7 – Etapes de la méthode Bootstrap

Les charges ultimes stochastiques vues en $N+1$ calculées lors de la dernière étape permettent de calculer le boni-mali (CDR) pour chaque simulation. La volatilité à un an correspond alors à l'écart-type de l'ensemble des boni-mali.

L'approche de la méthode de Merz-Wuthrich par procédure Bootstrap est détaillé dans l'article « One-year reserve risk including a tail factor : closed formula and Bootstrap approaches »

L'ensemble de ces méthodes et modèles reposent sur de nombreuses hypothèses qu'il est nécessaire de vérifier à travers des tests d'hypothèses. Dans le cadre de notre étude, nous nous focaliserons sur l'hypothèse d'indépendance entre années de survenance.

4. Tests des hypothèses de Merz-Wuthrich Chain-Ladder

a) Hypothèse d'indépendance entre années de survenance (H1)

Thomas Mack propose un test d'absence d'effet calendaire permettant de tester l'indépendance entre années de survenance. En effet, la présence d'effets calendaires due à un changement de gestion de sinistres, de provisionnement ou encore l'inflation peut remettre en cause cette hypothèse d'indépendance.

Ce test repose sur les diagonales du d-triangle, triangle des facteurs de développement individuel $f_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}$. Nous notons les éléments de la j -ème diagonale, $1 \leq j \leq N - 1$, de la manière suivante :

$$D_j = \{C_{j,1}, C_{j-1,2}, \dots, C_{1,j}\}$$

Les diagonales du d-triangle LR_j et LR_{j-1} s'expriment également en fonction des montants de charges (ou paiements) cumulés de la diagonale D_j :

$$LR_j = \left\{ \frac{C_{j,2}}{C_{j,1}}, \frac{C_{j-1,3}}{C_{j-1,2}}, \dots, \frac{C_{1,j+1}}{C_{1,j}} \right\} = \{f_{j,1}, f_{j-1,2}, \dots, f_{1,j}\}$$

$$LR_{j-1} = \left\{ \frac{C_{j-1,2}}{C_{j-1,1}}, \frac{C_{j-2,3}}{C_{j-2,2}}, \dots, \frac{C_{1,j}}{C_{1,j-1}} \right\} = \{f_{j-1,1}, f_{j-2,2}, \dots, f_{1,j-1}\}$$

Par conséquent, les diagonales LR_j et LR_{j-1} subissent les effets calendaires portés par D_j : une hausse (baisse) des éléments de D_j entraîne une hausse (baisse) des éléments de LR_{j-1} et une baisse (hausse) des éléments de LR_j .

Le principe du test est de séparer le d-triangle en répartissant les coefficients de développement de chaque colonne en deux groupes : les coefficients supérieurs à la médiane de la colonne et les coefficients inférieurs à la médiane de la colonne. A l'issue de cette répartition, nous regardons par diagonale quel est le groupe dominant. Nous notons pour la suite :

$$med_j = \text{médiane}(C_{1,j}, \dots, C_{n-j,j}), 1 \leq j \leq N - 1.$$

Les étapes du test sont les suivantes :

1. Nous construisons le triangle nouveau triangle issu de la séparation des coefficients de développement à savoir, pour une année de développement j fixée :

- Si $f_{i,j} < med_j$, le coefficient est marqué par un P
- Si $f_{i,j} > med_j$, le coefficient est marqué par un G

Les coefficients égaux à la médiane sont éliminés (cas des années de développement composées d'un nombre impair de coefficients).

2. Pour chaque diagonale LR_j comportant au moins 2 coefficients marqués, nous comptons le nombre de P et de G que nous notons respectivement P_j et G_j . Sous l'hypothèse d'absence d'effets calendaires, la probabilité d'appartenir à un des deux groupes de coefficients est égale à 50% et donc P_j et G_j devraient être proches.

3. Nous définissons la variable $Z_j = \min(P_j; G_j)$. D'après le point précédent, Z_j devrait être proche de $\frac{G_j + P_j}{2}$.

4. Sous l'hypothèse d'absence d'effets calendaires, les variables P_j et G_j suivent une loi binomiale $\mathcal{B}(N_j, p)$ avec $N_j = P_j + G_j$ et $p = 0,5$. Thomas Mack a démontré que l'espérance et la variance de cette variable aléatoire s'écrivent :

$$E(Z_j) = \frac{N_j}{2} - \binom{N_j - 1}{m_j} \frac{N_j}{2^{N_j}} \quad \text{avec} \quad \binom{n}{k} = \frac{n!}{k!(n-k)!} \quad \text{et} \quad m_j = \frac{N_j - 1}{2}$$

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

5. Nous définissons la variable aléatoire $Z = \sum_j Z_j$ qui est une variable aléatoire globale. Sous l'hypothèse d'absence d'effets calendaires, nous pouvons écrire :

$$E(Z) = \sum_j E(Z_j)$$

$$Var(Z) = \sum_j Var(Z_j)$$

6. Z étant la somme de variables aléatoires supposées suivre des lois binomiales $\mathcal{B}(N_j, p)$, elle suit approximativement une loi normale.

Pour ne pas rejeter l'hypothèse au seuil $\alpha = 0,05$, on doit donc avoir :

$$-q_{1-\frac{\alpha}{2}} \leq \frac{Z - E(Z)}{\sqrt{Var(Z)}} \leq q_{1-\frac{\alpha}{2}}$$

$$E(Z) - q_{1-\frac{\alpha}{2}}\sqrt{\text{Var}(Z)} \leq Z \leq E(Z) + q_{1-\frac{\alpha}{2}}\sqrt{\text{Var}(Z)}$$

Où $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale centrée

b) *Hypothèse de linéarité (H2)*

L'hypothèse de linéarité suppose que, pour une année de survenance donnée i , il existe une relation linéaire, entre les montants cumulés $C_{i,j}$ et $C_{i,j+1}$. Pour tester cette hypothèse, nous effectuons un test graphique ainsi qu'un test statistique de régression.

Le test graphique consiste à représenter pour chaque année de développement j , l'ensemble des couples $(C_{i,j}, C_{i,j+1})$. Ces couples doivent être alignés sur une droite de pente f_j et passant par l'origine.

Le test statistique de régression sur les couples $(C_{i,j}, C_{i,j+1})$ permet de compléter l'analyse graphique effectuée à l'aide des différents coefficients R^2 obtenus. Le test consiste à évaluer la moyenne et le minimum des coefficient R^2 sur l'ensemble des couples $(C_{i,j}, C_{i,j+1})$. Pour rappel, la qualité d'une régression linéaire est d'autant plus élevée que le coefficient R^2 est proche de 1.

L'hypothèse de linéarité sera considérée comme satisfaite lorsque le coefficient R^2 moyen, évalué à partir de l'ensemble des développements, est proche de 1. Le coefficient R^2 minimum est généralement atteint sur les premiers délais de développement, qui sont naturellement les moins stables, en particulier sur les branches longues. Ce coefficient R^2 minimum doit également être proche de 1 pour que les premiers délais de développement puissent être considérés comme vérifiant l'hypothèse de linéarité. Par jugement d'expert, le seuil permettant d'accepter ou rejeter l'hypothèse d'indépendance est fixé à 0,7.

c) *Hypothèse de variance (H3)*

L'hypothèse de variance suppose que la variance du montant $C_{i,j+1}$ est proportionnelle au montant $C_{i,j}$. Pour tester cette hypothèse, nous effectuons un test graphique, un test statistique du ρ de Spearman sur le triangle dans son ensemble.

Le test graphique consiste à représenter les résidus normalisés en fonction des montants cumulés $C_{i,j}$. L'hypothèse est rejetée si les résidus normalisés présentent une structure non aléatoire. Les résidus normalisés s'expriment de la manière suivante :

$$r_{i,j} = \frac{C_{i,j+1} - f_j C_{i,j}}{\sqrt{C_{i,j}}}$$

Le test statistique du ρ de Spearman permet de tester la non-corrélation entre les résidus normalisés sur le triangle dans son ensemble par année de développement. Le nombre d'années de survénance, noté N , est supposé égal au nombre d'années de développement.

Test statistique du ρ de Spearman :

Ce test a pour but d'accepter ou rejeter l'hypothèse de non-corrélation entre les résidus sur le triangle dans son ensemble :

$$H_0 : \rho = 0$$

Où ρ est la statistique du test.

Pour chaque colonne j tel que $N - j \leq 4$, les étapes du test statistique du ρ de Spearman sont les suivantes :

1. Tri des résidus par ordre croissant pour chaque colonne j en notant $u_{i,j}$ le rang des résidus $r_{i,j}$ dans la colonne j .
2. Tri par ordre croissant chaque colonne j privée du résidu de la dernière année de survénance $r_{N-j,j}$ en notant $v_{i,j}$ le rang des résidus $r_{i,j}$ dans cette nouvelle colonne j .
3. Calcul du coefficient de corrélation $\rho_{j,j+1}$ entre la colonne des rang $u_{i,j+1}$ et la colonne des rangs $v_{i,j}$:

$$\rho_{j,j+1} = \frac{\sum_{i=1}^{N-j-1} (u_{i,j+1} - \bar{u}_{j+1})(v_{i,j} - \bar{v}_j)}{\sqrt{\sum_{i=1}^{N-j-1} (u_{i,j+1} - \bar{u}_{j+1})^2} \sqrt{\sum_{i=1}^{N-j-1} (v_{i,j} - \bar{v}_j)^2}}$$

Avec $\bar{u}_{j+1} = \frac{1}{N-j-1} \sum_{i=1}^{N-j-1} u_{i,j+1}$ et $\bar{v}_j = \frac{1}{N-j-1} \sum_{i=1}^{N-j-1} v_{i,j}$

4. Sous l'hypothèse de non-corrélation entre résidus des différentes colonnes, l'espérance et la variance sont égales à :

$$\begin{aligned} \mathbb{E}(\rho_{j,j+1}) &= 0 \\ \text{Var}(\rho_{j,j+1}) &= \frac{1}{N-j-2} \end{aligned}$$

5. Afin d'obtenir une moyenne pondérée des $\rho_{j,j+1}$ de variance minimale, les poids de pondération doivent être inversement proportionnels à la variance des $\rho_{j,j+1}$. Nous posons alors $\omega_j = \frac{N-j-2}{\sum_{j=1}^{N-4} N-j-2}$, ce qui permet d'avoir $\sum_{j=1}^{N-4} \omega_j = 1$. La moyenne pondérée des $\rho_{j,j+1}$ s'écrit comme suit :

$$\rho = \sum_{j=1}^{N-4} \omega_j \rho_{j,j+1} = \sum_{j=1}^{N-4} \frac{(N-j-2) \rho_{j,j+1}}{\sum_{j=1}^{N-4} N-j-2}$$

Sous l'hypothèse H_0 , l'espérance et la variance valent :

$$\begin{aligned} \mathbb{E}(\rho) &= 0 \\ \text{Var}(\rho) &= \frac{1}{\sum_{j=1}^{N-4} N - j - 2} \end{aligned}$$

Nous pouvons supposer que ρ suit approximativement une loi normale.

6. L'hypothèse H_0 ne sera pas rejetée si :

$$-q_{0,975} \leq \frac{\rho}{\sqrt{\text{Var}(\rho)}} \leq q_{0,975}$$

Où $q_{0,975}$ est le quantile de la loi normale centrée réduite d'ordre $1 - \frac{\alpha}{2}$ avec $\alpha = 5\%$

Le calcul du SCR de risque de réserves de Groupama repose sur l'estimation de la volatilité à un an du modèle Merz-Wuthrich Chain-Ladder. Son application peut alors être remise en cause lorsqu'une des hypothèses est rejetée. Souvent, l'hypothèse d'indépendance entre années de survenance est la seule hypothèse rejetée parmi les trois hypothèses de la méthode Merz-Wuthrich Chain-Ladder. Afin de conserver le modèle actuel de calcul de volatilité à un an, il est donc nécessaire de trouver une méthode alternative au rejet de cette hypothèse.

Partie 2. THÉORIE DE LA
DÉPENDANCE ET DES MÉTHODES
ÉTUDIÉES

La méthode Merz-Wuthrich utilisée au sein du modèle interne partiel de Groupama présente des limites à travers ses hypothèses, notamment l'hypothèse d'indépendance entre années de survenance.

L'objectif de cette partie est de présenter l'ensemble des concepts et méthodes théoriques que nous appliquerons pour étudier la dépendance entre années de survenance. Dans un premier temps, nous présenterons l'une des potentielles causes du rejet du test d'indépendance entre années de survenance, en accord avec le contexte économique actuel : l'inflation. Ensuite, nous analyserons la dépendance entre années de survenance à travers différentes métriques. Pour finir, nous étudierons un modèle alternatif au modèle Chain-Ladder restant proche de ce dernier : le GLM Tweedie.

A. Prise en compte de l'inflation

Lors du paiement des sinistres, la compagnie d'assurance a préalablement estimé les coûts de l'indemnisation en tenant compte de nombreux facteurs variant en fonction du secteur touché par le sinistre : le prix de la main d'œuvre, les frais médicaux, les matières premières, etc. Cependant, les prix de ces paramètres peuvent évoluer d'une année à l'autre en raison de l'inflation. L'assureur doit donc prendre en compte cette variation probable de prix lors du calcul des provisions pour sinistres à payer (PSAP) pour éviter le sous-provisionnement en cas de hausse des prix ou alors le sur-provisionnement en cas de baisse des prix.

L'inflation étant évaluée d'une année à l'autre, elle peut créer un effet calendaire dans les triangles des données. Cet effet pourrait être à l'origine du rejet du test d'hypothèse d'indépendance entre années de survenance.

1. Méthode de retraitement de l'inflation

Dans certains cas, les montants peuvent subir des modifications afin d'avoir des résultats plus représentatifs de la réalité : un évènement atypique ou encore des années anciennes peuvent avoir un impact moindre sur la réalité actuelle.

La modification des triangles peut par exemple s'effectuer par produit matriciel entre le triangle et une matrice de valeurs $\Omega = (\omega_{i,j})_{(i,j) \in \{1, \dots, N\}^2}$. En notant T , le triangle de données brutes, le triangle modifié s'obtient par : $\tilde{T} = \Omega T$. Les éléments de ce nouveau triangle s'expriment comme suit : $\tilde{C}_{i,j} = \omega_{i,j} C_{i,j}$. Dans le modèle actuel, aucune modification n'est appliquée ($\omega_{i,j} = 1 \forall (i,j)$)

La mise en as-if du triangle peut être considérée comme un retraitement multiplicatif. En effet, pour obtenir un triangle hors inflation, la méthode revient à annuler les effets de l'inflation en le multipliant

par les inflations cumulées par année calendaire. Le retraitement de l'inflation se fait donc par ajout de facteurs multiplicatifs sur les diagonales du triangle. Les facteurs d'inflation s'expriment de la manière suivante :

$$\forall (i, j) \in \{1, \dots, N\}^2, \quad \omega_{i,j} = \prod_{k=1}^{i+j} (1 + \alpha_k)$$

Où α_k est le taux d'inflation de l'année k .

2. Contexte économique actuel

Compte tenu du contexte actuel de forte inflation, la méthode du Bootstrap à un an est remise en cause. En effet, l'utilisation de cette méthode implique l'utilisation des coefficients de Chain-Ladder f_j qui projettent une inflation historique et constante lors de l'estimation du triangle inférieur. Les facteurs de développement ne tiennent pas compte de la hausse d'inflation et pourrait mener à une sous-estimation de la charge ultime et des provisions. L'utilisation de la méthode Chain-Ladder pose donc problème au sein des modèles internes. Il est alors nécessaire de procéder à une refonte de la modélisation de l'inflation future ainsi que la méthode de calcul du SCR.

Une méthode de provisionnement déterministe permet de tenir compte de l'inflation en l'utilisant comme paramètre du modèle : la méthode de Séparation de Taylor présentée en annexe.

B. Concepts mathématiques de dépendance

1. Fonction de répartition jointe

Au cours des études statistiques, il est essentiel de comprendre la liaison entre les variables étudiées afin d'obtenir le meilleur résultat d'une analyse. La non-prise en compte de cette relation peut mener à des résultats erronés suite aux mauvaises analyses effectuées.

Plusieurs situations nécessitent la prise en compte de cette relation de dépendance comme en tarification, en gestion des risques ou encore lors du calcul du SCR.

Pour identifier cette relation, les notions de fonction de la dépendance entre variables et de fonction de répartition jointe ont été introduites.

Nous nous restreignons au cas de deux variables pour simplifier les notations.

Fonction de répartition jointe :

Etant donné un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, la fonction de répartition jointe des variables aléatoires X et Y est la fonction de répartition de la variable aléatoire (X, Y) . Elle est définie de la manière suivante :

$$\forall (a, b) \in \mathbb{R}, F_{(X,Y)}(a, b) = \mathbb{P}(X \leq a, Y \leq b)$$

Fonction de répartition individuelle :

Etant donné un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, la fonction de répartition individuelle d'une variable aléatoire X est définie de la manière suivante :

$$\forall a \in \mathbb{R}, F_X(a) = \mathbb{P}(X \leq a)$$

La fonction de répartition individuelle peut s'obtenir à l'aide de la fonction de répartition jointe :

$$F_X(a) = \mathbb{P}(X \leq a, Y \leq \infty)$$

$$F_X(a) = \mathbb{P}\left(\lim_{b \rightarrow \infty} \{X \leq a, Y \leq b\}\right)$$

$$F_X(a) = \lim_{b \rightarrow \infty} \mathbb{P}(\{X \leq a, Y \leq b\})$$

$$F_X(a) = \lim_{b \rightarrow \infty} F_{(X,Y)}(a, b)$$

$$F_X(a) = F_{(X,Y)}(a, \infty)$$

Cependant, la connaissance des lois individuelles ne permet de déterminer la loi jointe. Pour construire la loi jointe à partir des lois individuelles, il est nécessaire de connaître la structure de dépendance des variables.

2. Notion et mesures de dépendance

Considérons un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, deux variables aléatoires X et Y .

Indépendance :

Les variables aléatoires X et Y sont indépendantes si et seulement si la loi jointe de la variable aléatoire (X, Y) est le produit des lois individuelles de X et Y :

$$X \text{ et } Y \text{ indépendantes} \Leftrightarrow \forall (a, b) \in \mathbb{R}, F_{(X,Y)}(a, b) = F_X(a)F_Y(b)$$

Dépendance :

Les variables aléatoires X et Y sont dépendantes si et seulement si la loi jointe de la variable aléatoire (X, Y) est différente du produit des lois individuelles de X et Y :

$$X \text{ et } Y \text{ dépendantes} \Leftrightarrow \forall (a, b) \in \mathbb{R}, F_{(X,Y)}(a, b) \neq F_X(a)F_Y(b)$$

Les mesures de dépendance sont des fonctions permettent d'évaluer et quantifier la dépendance liant deux variables aléatoires. Cette définition est générale mais en pratique, une mesure de dépendance, notée δ , doit posséder les propriétés suivantes :

- Symétrie : $\delta(X, Y) = \delta(Y, X)$
- Normalisation : $-1 \leq \delta(X, Y) \leq 1$
- $\delta(X, Y) = 1 \Leftrightarrow (X, Y) =_{LOI} (F_X^{-1}(U), F_Y^{-1}(U))$ où $U \sim \mathbb{U}[0; 1]$
- $\delta(X, Y) = -1 \Leftrightarrow (X, Y) =_{LOI} (F_X^{-1}(U), F_Y^{-1}(1 - U))$ où $U \sim \mathbb{U}[0; 1]$
- Soit f une fonction strictement croissante, alors :

$$\delta(f(X), Y) = \begin{cases} \delta(X, Y), & f \text{ croissante} \\ -\delta(X, Y), & f \text{ décroissante} \end{cases}$$

En général l'équivalence suivante n'est pas vérifiée lorsque δ est une mesure de concordance :

$$X \text{ et } Y \text{ dépendantes} \Leftrightarrow \delta(X, Y) = 0$$

La plupart des mesures de dépendance rencontrées vérifient uniquement l'implication, à savoir :

$$X \text{ et } Y \text{ dépendantes} \Rightarrow \delta(X, Y) = 0$$

Nous allons présenter par la suite les mesures de dépendance les plus fréquentes.

a) *Corrélation de Pearson*

Le coefficient de corrélation de Pearson, aussi appelé coefficient de corrélation linéaire, mesure la qualité de l'approximation de Y par une fonction affine de X . Il s'exprime de la manière suivante :

$$\rho_{pearson} \stackrel{\text{def}}{=} \frac{COV(X, Y)}{\sigma_X \sigma_Y}$$

Où :

- $COV(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$: covariance entre X et Y .
- σ_X : écart-type de X
- σ_Y : écart-type de Y

Ce coefficient possède les propriétés suivantes :

- Symétrie : $\rho_{pearson}(X, Y) = \rho_{pearson}(Y, X)$

- Normalisation : $-1 \leq \rho_{pearson}(X, Y) \leq 1$
- Si $\rho_{pearson}(X, Y) = 1$ alors il y a corrélation positive parfaite entre X et Y . Il existe donc $a > 0$ et $b \in \mathbb{R}$ tels que : $Y = aX + b$. Il s'agit de variables co-monotones.
- Si $\rho_{pearson}(X, Y) = -1$ alors il y a corrélation négative parfaite entre X et Y . Il existe donc $a < 0$ et $b \in \mathbb{R}$ tels que : $Y = aX + b$. Il s'agit de variables anti-monotones.
- X et Y indépendantes $\Rightarrow \rho_{pearson}(X, Y) = 0$.

La corrélation de Pearson permet uniquement de mesurer le degré de liaison linéaire. Par conséquent, ce coefficient peut être nul tandis que les variables X et Y sont dépendantes. Ce coefficient ne respectant qu'une partie des propriétés des mesures de concordance, il ne peut être considéré comme une mesure de dépendance.

b) Corrélation des rangs

La notion de corrélation des rangs repose sur les concepts de concordance et de discordance. Considérons (X_1, Y_1) et (X_2, Y_2) deux couples de variables aléatoires

Concordance :

Les couples (X_1, Y_1) et (X_2, Y_2) sont en concordance si :

$$(X_1 - X_2)(Y_1 - Y_2) > 0 \Leftrightarrow (X_1 < X_2 \text{ et } Y_1 < Y_2) \text{ ou } (X_1 > X_2 \text{ et } Y_1 > Y_2)$$

Discordance :

Le couples (X_1, Y_1) et (X_2, Y_2) sont en discordance si :

$$(X_1 - X_2)(Y_1 - Y_2) < 0 \Leftrightarrow (X_1 < X_2 \text{ et } Y_1 > Y_2) \text{ ou } (X_1 > X_2 \text{ et } Y_1 < Y_2)$$

(1) Corrélation des rangs de Spearman

Le coefficient de corrélation des rangs de Spearman, noté $\rho_{spearman}$, est calculé sur les rangs des variables (X, Y) . En notant $(x_i, y_i)_{i \in \{1, \dots, N\}}$ un échantillon de (X, Y) et $(r_i, s_i)_{i \in \{1, \dots, N\}}$ les rangs correspondants, le coefficient de corrélation de Spearman se calcule de la manière suivante :

$$\rho_{spearman} \stackrel{\text{def}}{=} 1 - \frac{6 \sum_{i=1}^N (r_i - s_i)^2}{N(N^2 - 1)}$$

Il peut également s'exprimer à l'aide du coefficient de corrélation de Pearson :

$$\rho_{spearman} \stackrel{\text{def}}{=} \rho_{pearson}(F_X(X), F_Y(Y))$$

S'exprimant à l'aide du coefficient de corrélation de Pearson, le coefficient de corrélation des rangs de Spearman ne peut pas être considéré comme une mesure de dépendance.

(2) Corrélation des rangs de Kendall

Le coefficient de corrélation des rangs de Kendall, noté $\tau_{kendall}$, est calculé comme suit :

$$\tau_{kendall} \stackrel{\text{def}}{=} \mathbb{E} \left(\text{signe} \left((X - \tilde{X})(Y - \tilde{Y}) \right) \right)$$

Où (\tilde{X}, \tilde{Y}) est un couple indépendant de (X, Y) et identiquement distribué.

Le coefficient de corrélation des rangs de Kendall est une mesure de concordance. Il compare la probabilité que deux couples indépendants et identiquement distribués soient en concordance avec la probabilité qu'ils soient en discordance.

Il s'exprime de la manière suivante :

$$\tau_{kendall} \stackrel{\text{def}}{=} \mathbb{P} \left((X - \tilde{X})(Y - \tilde{Y}) > 0 \right) - \mathbb{P} \left((X - \tilde{X})(Y - \tilde{Y}) < 0 \right)$$

Les mesures présentées permettent de mesurer le niveau de dépendance existant entre deux variables. La théorie des copules quant à elle, permet d'extraire une forme et un niveau de dépendance entre variables aléatoires.

3. Théorie des copules

La distribution jointe décrit entièrement la dépendance entre variables aléatoires. Cependant, il est possible de distinguer les comportements des distributions marginales de la structure de dépendance.

La copule est l'outil statistique permettant d'explicitier la structure de dépendance d'une distribution jointe afin d'identifier et séparer les distributions marginales et la dépendance.

La théorie des copules multivariées est une extension du cas bivarié. Comme précédemment, nous nous limitons au cas bivarié afin de simplifier les écritures.

a) *Définition et caractérisation*

Les données des lois marginales de X et Y ne déterminent pas la loi du vecteur aléatoire (X, Y) , i.e. la loi jointe de X et Y . Les copules permettent une modélisation séparée des lois marginales de X et Y et de la structure de dépendance entre X et Y exprimée par une copule :

$$\text{Loi}(X), \text{Loi}(Y) \xrightarrow{\text{copule}} \text{Loi}(X, Y)$$

Définition :

Une copule de dimension $d = 2$ est définie comme la fonction de répartition d'un vecteur aléatoire (U_1, U_2) où $U_i \sim \mathbb{U}[0; 1], i \in \{1, 2\}$. En notant C la copule bivariée, nous pouvons écrire :

$$C : [0, 1]^d \rightarrow [0, 1]$$

$$C(u_1, u_2) = \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2)$$

Caractérisation :

Une fonction $C : [0, 1]^d \rightarrow [0, 1]$ est une copule bivariée si et seulement si elle possède les propriétés suivantes :

- $\forall (u_1, u_2) \in [0, 1], C(u_1, 0) = C(0, u_2) = 0$
- $\forall (u_1, u_2) \in [0, 1], C(u_1, 1) = u_1$ et $C(1, u_2) = u_2$
- C est croissante : $\forall (u_1, u_2, v_1, v_2) \in [0, 1]^4$ tels que $u_1 < u_2$ et $v_1 < v_2$,
 $C(u_2, v_2) - C(u_2, v_1) - (C(u_1, v_2) - C(u_1, v_1)) \geq 0$

b) *Théorème de Sklar*

Etant donné un vecteur aléatoire (X, Y) de fonction de répartition $F_{X, Y}$, il existe une copule C telle que :

$$\forall (x, y) \in \mathbb{R}, \quad F_{X, Y}(x, y) = C(F_X(x), F_Y(y))$$

La loi du vecteur (X, Y) est entièrement déterminée par les lois marginales F_X et F_Y lorsqu'il admet C pour copule. Si les lois individuelles de X et Y sont continues, alors la copule C est unique. Dans le cas de lois marginales non continues, il est toujours possible de définir une copule mais elle n'est plus unique. L'utilisation de copule perd alors de l'intérêt : la structure de dépendance n'est plus unique.

Réciproquement, étant donnée une copule C et des fonctions de répartition individuelles F_X et F_Y , il existe toujours un vecteur aléatoire (X, Y) dont les lois marginales sont données par F_X et F_Y et admettant C pour copule. Ainsi, il est possible de déterminer la loi jointe du vecteur aléatoire :

$$\forall (x, y) \in \mathbb{R}, \quad F_{X,Y}(x, y) = C(F_X(x), F_Y(y))$$

Dans la suite, nous supposons que les lois marginales X et Y sont continues.

c) Propriétés

Copule de survie :

Si (X, Y) est un vecteur aléatoire de copule C , nous pouvons écrire :

$$\forall (x, y) \in \mathbb{R}, \quad \mathbb{P}(X > x, Y > y) = C^*(\mathbb{P}(X > x), \mathbb{P}(Y > y))$$

Où C^* est la copule de survie du vecteur (X, Y) . C'est la fonction de répartition du vecteur aléatoire $(1 - F_X(X), 1 - F_Y(Y))$, ou encore celle du vecteur $(1 - U_1, 1 - U_2)$, où $U_i \sim \mathbb{U}[0; 1], i \in \{1, 2\}$ de fonction de répartition C .

Invariance par transformation croissante :

Si (X, Y) est un vecteur aléatoire de copule C et si (φ_X, φ_Y) sont des fonctions strictement croissantes, alors le vecteur défini par $(\varphi_X(X), \varphi_Y(Y))$ admet également C pour copule. Un changement d'échelle sur les variables aléatoires ne modifie pas la copule.

Densité :

En cas d'existence, la densité d'une copule bivariée est définie par :

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v}$$

Si (X, Y) est un vecteur aléatoire de copule C de densité c et si les variables aléatoires X et Y possèdent des densités f_X et f_Y sur \mathbb{R} , alors la densité du vecteur aléatoire (X, Y) est définie par :

$$\forall (x, y) \in \mathbb{R}, \quad f_{X,Y}(x, y) = c(F_X(x), F_Y(y)) \times f_X(x) \times f_Y(y)$$

d) Exemples de base

Copule d'indépendance :

Elle correspond au cas où les variables aléatoires X et Y sont indépendantes :

$$\forall (u, v) \in [0, 1]^2, \quad C(u, v) \stackrel{\text{def}}{=} u \times v$$

La copule C est la fonction de répartition du vecteur (U_1, U_2) où $U_i \sim \mathbb{U}[0; 1], i \in \{1, 2\}$.

Copule co-monotone :

Elle correspond au cas de dépendance positive parfaite entre les variables aléatoires X et Y :

$$\forall (u, v) \in [0, 1]^2, \quad C(u, v) \stackrel{\text{def}}{=} \min(u, v)$$

Copule anti-monotone :

Elle correspond au cas de dépendance négative parfaite entre les variables aléatoires X et Y :

$$\forall (u, v) \in [0, 1]^2, \quad C(u, v) \stackrel{\text{def}}{=} \max(u + v - 1, 0)$$

Copules elliptiques :

Cette famille de copules est adaptée aux distributions elliptiques multivariées. La famille des distributions elliptiques englobe l'ensemble des distributions généralisant la distribution normale multivariée.

La famille des copules elliptiques comprend notamment les copules suivantes :

- Copule normale
- Copule de t-Student

Copule normale

C'est la copule d'un vecteur gaussien bi-dimensionnel de matrice de variance-covariance Σ :

$$\forall (u_1, u_2) \in [0, 1]^2, \quad C(u_1, u_2) = \mathbb{P}(F_{Y_1}(Y_1) \leq u_1, F_{Y_2}(Y_2) \leq u_2)$$

Où :

- $(Y_1, Y_2) \sim \mathcal{N}(0, \Sigma)$
- $\Sigma = \begin{pmatrix} \sigma_1^2 & \alpha\sigma_1\sigma_2 \\ \alpha\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ où α : coefficient de corrélation entre Y_1 et Y_2
- F_{Y_i} : fonction de répartition de la loi $\mathcal{N}(0, \sigma_i^2), i \in \{1, 2\}$

En notant F_{Y_1, Y_2} la fonction de répartition du vecteur (Y_1, Y_2) , elle s'exprime également de la manière suivante, $\forall (u_1, u_2) \in [0, 1]^2$:

$$C(u_1, u_2) = F_{Y_1, Y_2}(F_{Y_1}^{-1}(u_1), F_{Y_2}^{-1}(u_2))$$

$$C(u_1, u_2) = \frac{1}{2\pi\sqrt{1-\alpha^2}} \int_{-\infty}^{F_{Y_1}^{-1}(u_1)} \int_{-\infty}^{F_{Y_2}^{-1}(u_2)} \exp\left(\frac{-(x^2 + y^2 - 2\alpha xy)}{2(1-\alpha^2)}\right) dx dy$$

Copule de t-Student

C'est la copule d'un vecteur aléatoire de loi de Student bi-dimensionnel de matrice de variance-covariance Σ .

La fonction de répartition de la loi de Student est définie de la manière suivante :

$$\forall x \in \mathbb{R}, \quad t_\nu(x) = \int_{-\infty}^x \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu} \Gamma(\frac{\nu}{2})} \left(1 + \frac{s^2}{\nu}\right)^{-\frac{\nu+1}{2}} ds$$

Où $\Gamma(\cdot)$ est la fonction Gamma.

La fonction de répartition de la loi de Student bivariée est définie comme suit :

$$\forall (x, y) \in \mathbb{R}^2, \quad t_{\nu, \alpha}(x, y) = \int_{-\infty}^x \int_{-\infty}^y \frac{1}{2\pi\sqrt{1-\alpha^2}} \left(1 + \frac{s^2 + r^2 - 2\alpha sr}{\nu(1-\alpha^2)}\right)^{-\frac{\nu+2}{2}} ds dr$$

Où α est le paramètre de dépendance, $\alpha \in]0, 1[$

La copule de Student s'écrit donc :

$$\forall (u_1, u_2) \in [0, 1]^2, \quad C(u_1, u_2) = t_{\nu, \alpha}(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2))$$

Copules archimédiennes :

Les copules de cette famille sont de la forme :

$$\forall (u_1, u_2) \in [0, 1]^2, \quad C(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2))$$

Où $\phi :]0, 1] \rightarrow [0, +\infty[$ est une fonction appelée générateur de la copule. Pour que C soit une copule, le générateur doit posséder les propriétés suivantes :

- $\phi :]0, 1] \rightarrow [0, +\infty[$ est une bijection continue décroissante
- $\phi(1) = 0$ et $\phi(0) \rightarrow +\infty$
- ϕ est monotone

La famille des copules archimédiennes comprend notamment les copules suivantes :

- Copule de Clayton : $\phi(t) = \frac{1}{\theta}(t^{-\theta} - 1), \theta \geq 0$
- Copule de Frank : $\phi(t) = \log\left(\frac{e^{-t\theta} - 1}{e^{-\theta} - 1}\right), \theta \in \mathbb{R}$
- Copule de Gumbel : $\phi(t) = (-\log(t))^\theta, \theta \geq 1$

C. Modèles linéaires généralisés (GLM)

1. Définition

L'objectif d'une régression est de modéliser la valeur d'une variable y en fonction des valeurs de n variables (x_1, \dots, x_n) . La variable y est la variable réponse (expliquée) et les variables (x_1, \dots, x_n) sont les co-variables (explicatives). Dans une compagnie d'assurance, y peut par exemple représenter le coût d'un sinistre et (x_1, \dots, x_n) les caractéristiques de la police (âge de l'assuré, zone géographique, type de véhicule, ...). En général, les valeurs de y et de $x = (x_1, \dots, x_n)$ sont modélisées à l'aide d'un couple de variables aléatoires (Y, X) où $X = (X_1, \dots, X_n)$.

Les modèles linéaires généralisés est une méthode de régression, généralisant les modèles linéaires. Ils sont constitués de trois composantes essentielles :

- Une composante systématique : c'est une fonction affine des co-variables (ou d'un encodage) $x = (x_1, \dots, x_n)$. Elle représente la partie explicative du modèle et est notée η . Elle est définie par :

$$\eta = \sum_{i=0}^n \beta_i x_i$$

Et peut s'écrire de manière matricielle :

$$\eta = XB$$

Où $B = (\beta_1, \dots, \beta_n)^t$ est le vecteur des paramètres du modèle. Ce vecteur peut contenir un terme constant (intercept) et devient alors : $B = (\alpha, \beta_1, \dots, \beta_n)^t$.

- Une composante aléatoire : Elle spécifie le type de loi de la variable aléatoire $Y|X = x$ au sein de la famille exponentielle. Une loi appartient à la famille exponentielle si elle possède une densité de la forme :

$$f(y, \mu, \phi) = \exp\left(\frac{y\theta - b(\theta)}{\phi} + c(y, \phi)\right)$$

Où :

- o θ est le paramètre naturel
 - o $\phi > 0$ est le paramètre de dispersion
 - o $b : \mathbb{R} \rightarrow \mathbb{R}$ deux fois dérivable
 - o $c : \mathbb{R}^2 \rightarrow \mathbb{R}^2$
- Une fonction de lien : Notée g , elle spécifie la relation entre les composantes aléatoires et systématiques du modèle. Elle est définie comme suit, en notant $\mu = \mathbb{E}(Y|X = x)$:

$$g(\mu) = \eta$$

Elle est différentiable, continue et inversible.

L'espérance et la variance des lois de famille exponentielle sont définies par :

- $\mathbb{E}(Y) = b'(\theta)$
- $\mathbb{V}(Y) = b''(\theta)\phi$

Il est possible de reparamétriser le modèle en fonction de $\mu = \mathbb{E}(Y)$ en posant :

- $\theta = (b')^{-1}(\mu)$
- $\mathbb{V}(Y) = \phi v(\mu)$, où $v(\mu) \stackrel{\text{def}}{=} b''((b')^{-1}(\mu))$ est la fonction de variance.

La famille exponentielle englobe de nombreuses lois classiques telles que :

- La loi Normale
- La loi de Poisson
- La loi Gamma
- La loi Binomiale

Mais aussi d'autres lois moins classiques comme :

- La loi Binomiale Négative
- La loi Gaussienne Inverse
- La loi Tweedie

Dans notre étude, les modèles linéaires généralisés sont utilisés afin de prédire le montant des provisions en cherchant la distribution de probabilité qui correspond le mieux à la distribution des provisions. Nous nous sommes intéressés en particulier aux modèles Poisson et Tweedie.

Loi	Notation	Fonction de lien	$v(\mu)$
Poisson	$\mathcal{P}(\lambda)$	$\ln(\mu)$	μ
Tweedie	$\mathcal{PC}(\lambda, \Gamma(k, \alpha))$ $p \in]1, 2[$	$\frac{-\mu^{-(p-1)}}{p-1}$	μ^p

Figure 8 – Caractéristiques des GLM étudiés

Le terme de lois Tweedie désigne les lois de la famille exponentielle dont la fonction de variance est une puissance de l'espérance. Des valeurs particulières de p correspondent à des lois connues :

- $p = 0$: loi Normale
- $p = 1$: loi de Poisson
- $p = 2$: loi Gamma
- $p = 3$: loi Gaussienne Inverse

Dans le cas présenté dans le tableau ci-dessus, le paramètre de variance p est compris dans l'intervalle $]1,2[$ ce qui correspond au cas d'une loi de Poisson composée avec une loi Gamma. Dans la suite du mémoire, nous utiliserons le terme « Tweedie » pour désigner ce type de loi.

2. Provisionnement et GLM

Notations :

- $Y_{i,j}$: paiement incrémental d'année de survenance i et de développement j
- c_j : paramètre correspondant aux effets de la colonne j
- l_i : paramètre correspondant aux effets de la colonne i

L'utilisation des GLM en provisionnement consiste à supposer que les paiements incrémentaux $Y_{i,j}$ peuvent être prédits par une variable aléatoire de loi \mathcal{L} , dépendante de deux variables : une variable liée aux années de développement j et une variable liée aux années de survenance i :

$$Y_{i,j} \sim \mathcal{L}(c_j, l_i)$$

L'estimation des réserves par années de survenance i , notée \hat{R}_i , s'obtient par la relation suivante :

$$\hat{R}_i = \sum_{k=N-i+1}^N \hat{Y}_{i,k}$$

La réserve totale correspond à la somme des réserves individuelles :

$$\hat{R} = \sum_{i=1}^N \hat{R}_i$$

3. Procédure Bootstrap

La méthode Bootstrap est également applicable lors de l'utilisation d'un modèle GLM. Les étapes sont les suivantes :

1. Estimation des paramètres du modèle à partir du triangle initial des paiements incrémentaux $(Y_{i,j})_{i+j \leq N+1}$.
2. Calcul des résidus de Pearson :

$$r_{i,j} = \frac{Y_{i,j} - \hat{\mu}_{i,j}}{\sqrt{v(\hat{\mu}_{i,j})}}$$

Les étapes suivantes sont répétées B fois pour obtenir une distribution empirique de la charge ultime.

3. Rééchantillonnage des résidus de Pearson $r_{i,j}$ et construction du nouveau triangle des résidus $r_{i,j}^*$.
4. Calcul de la partie supérieur du triangle des paiements incrémentaux :

$$\forall i + j \leq N + 1, \quad Y_{i,j}^* = r_{i,j}^* \sqrt{v(\hat{\mu}_{i,j})} + \hat{\mu}_{i,j}$$

5. Application du modèle GLM sur ce nouveau triangle incrémental et détermination de la partie inférieure :

$$\forall i + j > N + 1, \quad Y_{i,j}^* = \mu_{i,j}^*$$

6. Calcul des provisions par années de survenance i , \hat{R}_i^* et du montant de la provision totale \hat{R}^* .

Après avoir présenté la théorie sous-jacente aux différentes méthodes que nous allons étudier, nous passons maintenant à l'application et l'étude de ces dernières. Pour ce faire, nous avons fait le choix d'appliquer les résultats sur les lignes d'activité medical expense et income protection. Ces deux lignes d'activité constituent un échantillon représentatif du risque porté par Groupama, ce qui a orienté notre choix vers ces dernières. La ligne d'activité medical expenses est une branche à développement court et la ligne d'activité income protection est une branche à développement long.

**Partie 3. APPLICATION ET
RÉSULTATS DES MÉTHODES
ÉTUDIÉES**

Cette partie est consacrée à la présentation des données ainsi qu'à l'analyse des résultats obtenus à travers les applications des différentes méthodes présentées sur les lignes d'activité medical expenses et income protection.

Dans un premier temps, nous nous pencherons sur l'inflation, un facteur d'actualité pouvant être la cause de la présence d'effet calendaire au sein du triangle des paiements. Ensuite, nous analyserons la corrélation existante entre année de survenance et étudierons l'impact d'un ajout de corrélation à travers deux méthodes. Pour finir, nous étudierons l'impact de l'utilisation d'un GLM Tweedie sur la volatilité à un an et par conséquent, sur la corrélation entre années de survenance au niveau des volatilité à un an individuelles.

A. Données étudiées

Pour mener notre étude, nous avons choisi les LoB suivantes :

- medical expenses (assurance frais médicaux) incluant l'ensemble des engagements couvrant les frais de soins.
- income protection (assurance protection du revenu) incluant l'ensemble des engagements couvrant la perte de revenus.

Nous pouvons distinguer deux types de branche en assurance non-vie : les branches courtes et les branches longues. De manière générale, une branche courte est définie comme une branche dont les paiements (ou charges) se stabilisent assez vite en fonction des années de développement. Une branche longue est, au contraire, caractérisée par un développement assez long des paiements (ou charges). En pratique, lorsqu'une branche est courte, c'est le triangle des paiements cumulés qui est utilisé ; lorsqu'une branche est longue, c'est le triangle des charges cumulés qui est utilisés. Il n'existe pas de règle de décision universelle permettant de caractériser les branches.

Suivant la gouvernance interne, la règle de décision retenue pour classer les branches consiste à vérifier que la durée des paiements de la LoB est inférieure à 3 ans (1^{er} critère) et que le nombre d'année de développement nécessaire pour atteindre 99,5% de l'ultime est inférieur à 5 ans (2nd critère).

1^{er} critère : il correspond à la durée des engagements pour chaque branche, calculée de la manière suivante :

$$duration = \sum_{j=1}^N e_j \times j$$

Où :

- $e_j = DM_j - DM_{j-1}$
- $DM_j = \frac{\sum_i C_{i,j}}{\sum_i U_i}$: développement moyen pondéré, calculés à partir de la deuxième année de développement.
- $C_{i,j}$: paiement (ou charge) d'année de survenance i et d'année de développement j
- U_i : charge ultime relative à l'année de survenance i

2nd critère : il correspond au nombre d'années de développement, noté m , nécessaire pour atteindre 99,5% de l'ultime. Cela revient à vérifier la condition suivante :

$$\left| \frac{\text{ultime} - \text{paiement}_m \text{ (ou charge}_m)}{\text{ultime}} \right| < 0,5\%$$

Les résultats des deux critères sont présentés ci-dessous :

LoB	Duration
med.exp.	0,53
inc.pr.	2,77

Figure 9 – Critère de duration (en année)

DY	ratio
1996	19,54%
1997	0,31%
1998	0,04%

Figure 10 – Critère de développement sur les paiements (medical expenses)

DY	ratio
1996	25,76%
1997	6,50%
1998	6,66%
1999	5,49%
2000	4,53%
2001	3,64%
2002	3,22%
2003	3,00%
2004	2,77%
2005	2,69%
2006	2,53%
2007	2,44%
2008	2,39%
2009	2,38%
2010	2,31%
2011	2,29%
2012	2,27%
2013	2,26%
2014	2,27%
2015	2,26%
2016	2,28%
2017	2,23%
2018	2,25%
2019	2,28%
2020	2,34%
2021	2,28%

Figure 12 – Critère de développement sur les paiements (income protection)

DY	ratio
1996	27,42%
1997	4,13%
1998	4,28%
1999	3,13%
2000	2,19%
2001	1,33%
2002	0,92%
2003	0,70%
2004	0,47%
2005	0,39%

Figure 11 – Critère de développement sur les charges (income protection)

Pour la LoB medical expenses, la duration est inférieure à 3 ans. De plus, le second critère montre que seulement 2 années de développement sont nécessaires pour passer sous le seuil de 0,5%, ce qui est inférieur à 5 ans. Nous pouvons donc définir la LoB medical expenses comme une branche courte et utiliser le triangle des paiements cumulés.

Pour la LoB income protection, la duration est inférieure à 3 ans mais l'écart est moindre (0,23 années). Le second critère appliqué sur les paiements n'est pas atteint même au bout de 26 années de développement, ce qui correspond à la profondeur de notre triangle. Il est donc nécessaire d'appliquer ce critère sur les charges. Ainsi, 9 années de développement sont nécessaires pour passer sous le seuil de 0,5%, ce qui est supérieur 5 ans. Compte tenu des résultats du second critère, nous pouvons qualifier la LoB income protection de branche longue et utiliser le triangle des charges cumulés.

Le nombre d'années de survenance et d'années de développement des deux triangles est identique et égal à 26. Les triangles utilisés sont présentés en annexes.

B. Résultats des tests d'hypothèse

1. Hypothèse d'indépendance entre années de survenance (H1)

Les résultats du test d'absence d'effets calendaires sont présentés ci-dessous :

LoB	Z	E(Z)	Var(Z)	Intervalle de confiance		Hypothèse
med.exp	69	78,5	19,5	69,8	87,1	Rejetée
inc.pr	106	122,5	28,4	112,0	132,9	Rejetée

Figure 13 – Résultats du test d'hypothèse d'indépendance

Pour la LoB medical expenses (med.exp.), la statistique de test vaut 69. Cette valeur n'appartient pas à l'intervalle de confiance [69,8 ; 87,1].

Pour la LoB income protection (inc.pr.), la statistique de test vaut 106. Cette valeur n'appartient pas à l'intervalle de confiance [112,0 ; 132,9].

2. Hypothèse de linéarité (H2)

Les résultats du test graphique et du test statistique de régression des deux lignes d'activité étudiées sont présentés ci-dessous. Sur chaque graphique nous retrouvons l'équation de régression de la droite ainsi que la valeur du R^2 de chaque couple $(C_{i,j}, C_{i,j+1})$.

Les premiers résultats présentés ci-dessous concernent la LoB medical expenses.

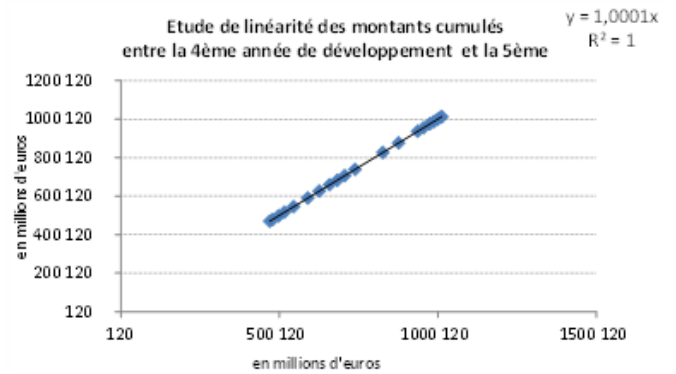
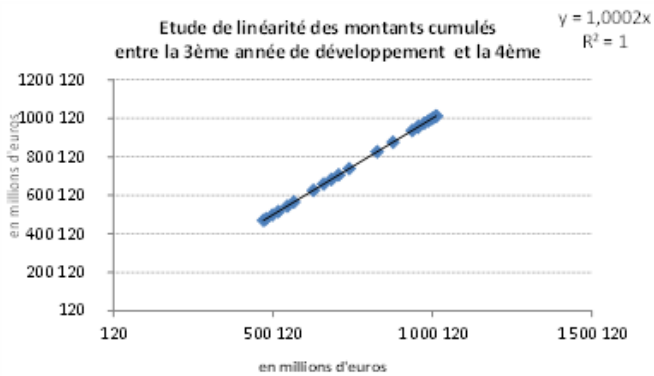
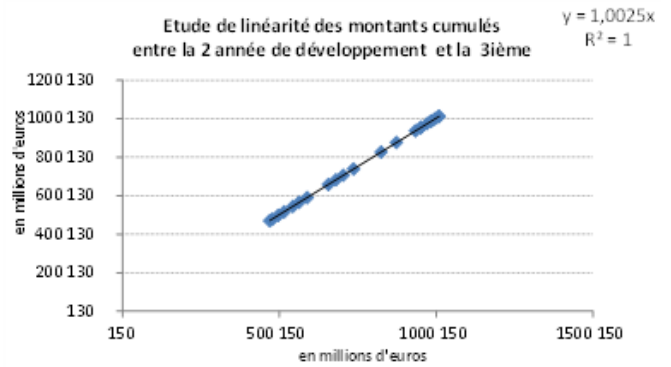
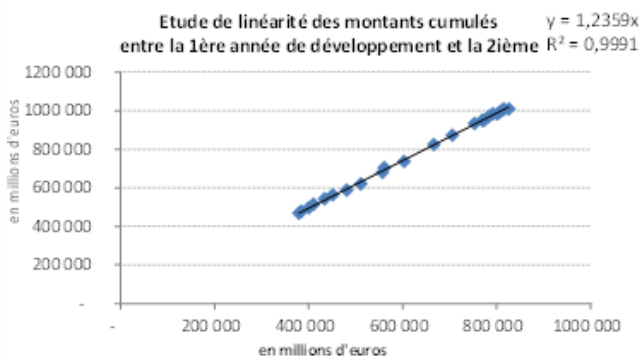


Figure 14 – Résultats du test graphique de l'hypothèse de linéarité des 4 premiers couples (medical expenses)

délagi de développement	facteur de développement moyen	penle de la droite
1	1,2365	1,2359
2	1,0024	1,0025
3	1,0002	1,0002
4	1,0001	1,0001

Figure 15 – Facteurs de développement et penles de régression des 4 premiers délagis de développement (medical expenses)

Nous avons fait le choix de montrer les résultats des quatre premiers délagis de développement car les résultats se stabilisent et deviennent identiques.

Le test graphique montre que les points de chaque couple $(C_{i,j}, C_{i,j+1})$ sont alignés selon une droite passant par l'origine. De plus, la penle de la droite correspond, à quelques millièmes près, aux facteurs de développement de Chain Ladder.

Le R^2 minimum vaut 0,991 et correspond au premier délagi de développement qui est souvent instable. Ce R^2 étant supérieur au seuil fixé de 0,7 permet de valider le test statistique de régression. L'ensemble

des R^2 des délais de développement suivants sont égaux à 1. La qualité de régression des couples $(C_{i,j}, C_{i,j+1})$ est parfaite.

Ainsi, les résultats du test graphique et du test statistique de régression permettent de valider l'hypothèse de linéarité entre les montants cumulés $C_{i,j}$ et $C_{i,j+1}$.

Les résultats suivants concernent la LoB income protection.

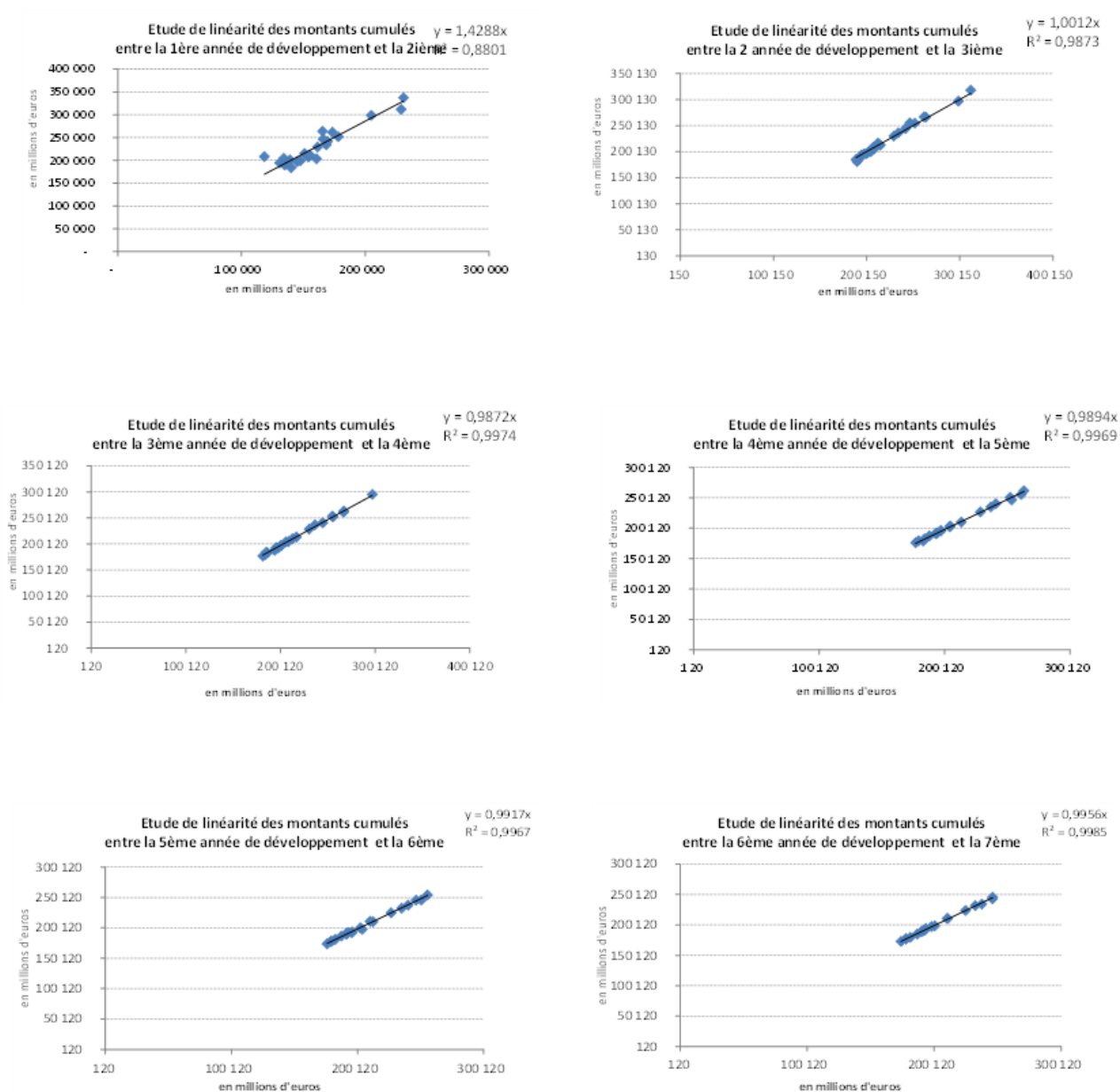


Figure 16 – Résultats du test graphique de l'hypothèse de linéarité des 6 premiers couples (income protection)

délagi de développement	facteur de développement moyen	penle de la droite
1	1,4312	1,4288
2	0,9993	1,0012
3	0,9868	0,9872
4	0,9896	0,9894
5	0,9918	0,9917
6	0,9957	0,9956

Figure 17 – Facteurs de développement et penles de régression des 4 premiers délags de développement (income protection)

Nous avons fait le choix de montrer les résultats des six premiers délags de développement car les résultats se stabilisent et tendent vers une régression parfaite ($R^2 = 1$).

Le test graphique montre que les points du premier développement, à savoir les points du couple $(C_{i,1}, C_{i,2})$ ne sont pas exactement alignés suivant la droite de régression. Cependant, le R^2 de la régression entre les points du couple vaut 0,8801, ce qui est supérieur à 0,7, seuil au-dessus duquel nous jugeons qu'une régression est de bonne qualité. De plus, la penle de la droite est très proche du facteur de développement de Chain Ladder (à 0,0024 près). Les résultats de ces deux tests combinés permettent d'accepter l'hypothèse de linéarité pour ce couple.

Pour chaque couple $(C_{i,j}, C_{i,j+1}), j > 1$, les points sont alignés selon une droite passant par l'origine. De plus, la penle de la droite correspond, à quelques millièmes près, aux facteurs de développement de Chain Ladder.

Le R^2 minimum vaut 0,8801 et correspond au premier délag de développement qui, comme nous l'avons vu, est instable. Ce R^2 étant supérieur au seuil fixé de 0,7 permet de valider le test statistique de régression. L'ensemble des R^2 des délags de développement suivants tendent vers 1. Le R^2 moyen de l'ensemble des développements vaut 0,9904, ce qui est largement supérieur à 0,7 et proche de 1. La régression des couples $(C_{i,j}, C_{i,j+1})$ est donc de bonne qualité.

Ainsi, les résultats du test graphique et du test statistique de régression permettent de valider l'hypothèse de linéarité entre les montants cumulés $C_{i,j}$ et $C_{i,j+1}$.

3. Hypothèse de variance (H3)

Les résultats du test graphique et du test statistique du ρ de Spearman des deux lignes d'activité étudiées sont présentés ci-dessous. Sur chaque graphique nous retrouvons le nuage de points des résidus en fonction des montants cumulés des années de développement.

Les premiers résultats présentés concernent la LoB medical expenses.

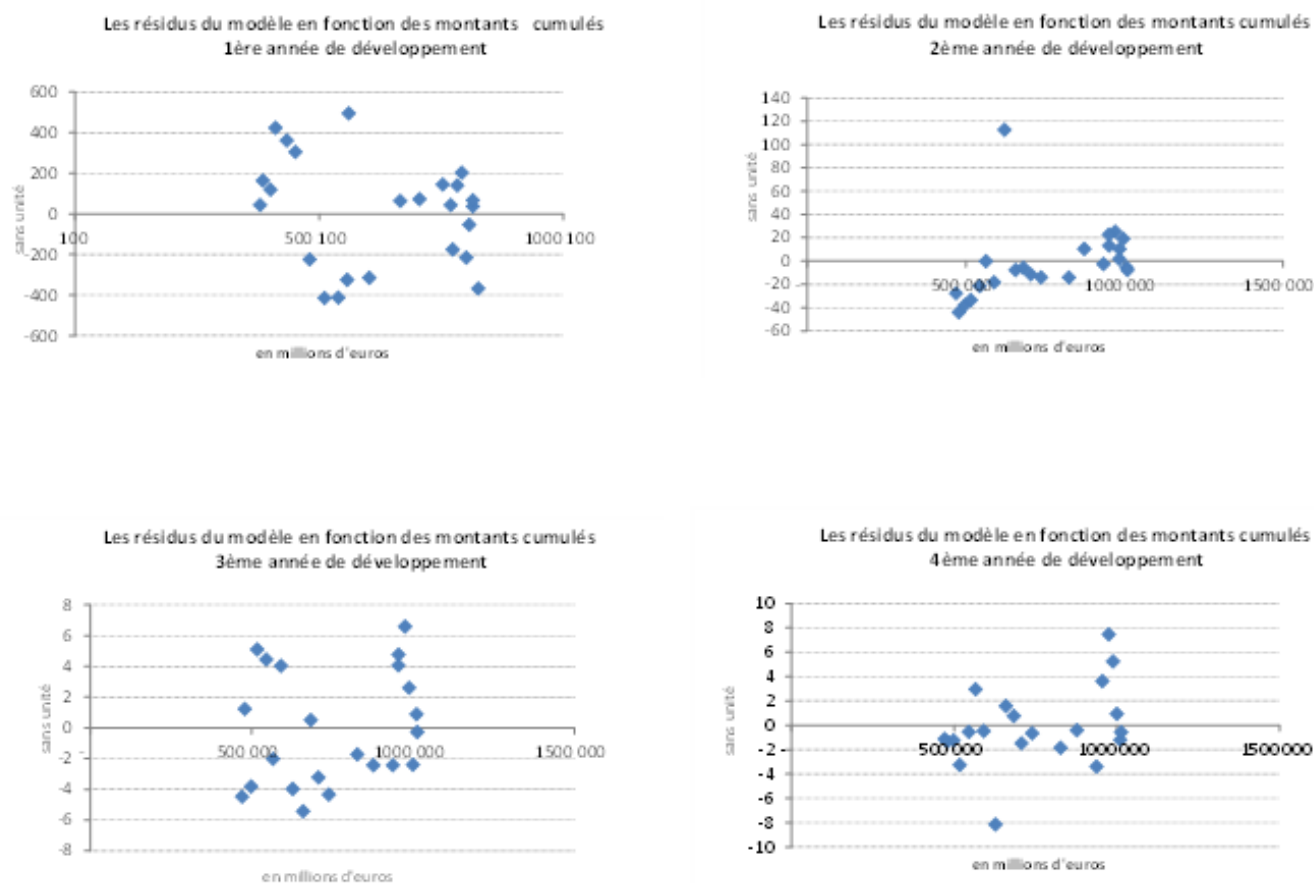


Figure 19 – Résultats du test graphique de l'hypothèse de variance des quatre premières années de développement (medical expenses)

LoB	T	E(T)	Var(T)	Intervalle à 95%		Hypothèse
med.exp	0,1457	0	0,0056	-0,1461	0,1461	Acceptée

Figure 18 – Résultats du test statistique de l'hypothèse de variance (medical expenses)

Nous avons fait le choix de montrer les résultats des tests graphiques des quatre premières années de développement car les nuages de points des années suivantes présentent une structure aléatoire similaire.

Les tests graphiques montrent que les résidus présentent une forme aléatoire à l'exception du nuage de points des résidus de la deuxième année de développement.

La statistique de test vaut 0,1457 et appartient à l'intervalle de confiance $[-0,1461; 0,1461]$. L'hypothèse H_0 de non-corrélation des résidus des résidus normalisés sur le triangle dans son ensemble par années de survenance est donc acceptée.

Les résultats du test graphique et du test statistique du ρ de Spearman permettent de valider l'hypothèse de variance pour la LoB medical expenses.

Les résultats suivants concernent la LoB income protection.

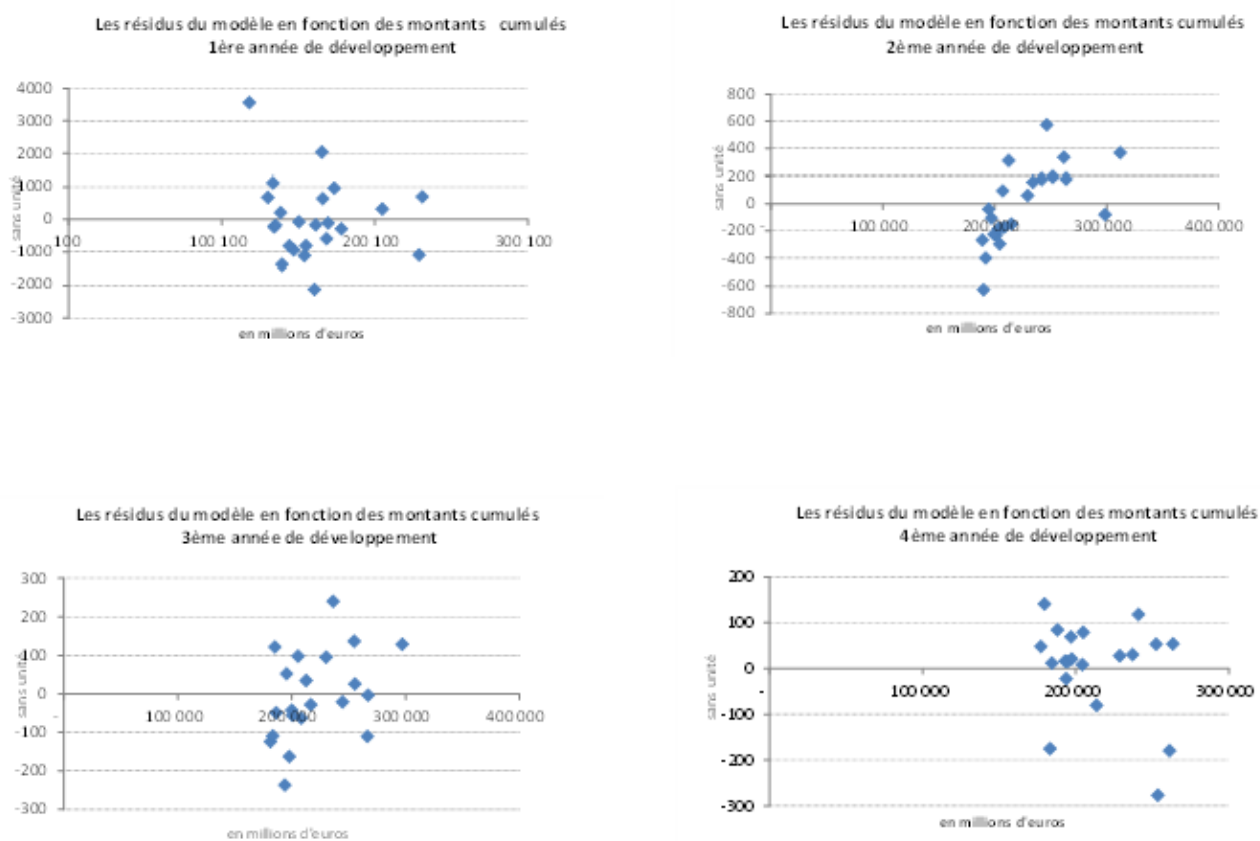


Figure 21 – Résultats du test graphique de l'hypothèse de variance des quatre premières années de développement (income protection)

LoB	T	E(T)	Var(T)	Intervalle à 95%		Hypothèse
inc.pr	0,0077	0	0,0052	-0,1407	0,1407	Acceptée

Figure 20 – Résultats du test statistique de l'hypothèse de variance (income protection)

Nous avons fait le choix de montrer les résultats des quatre premières années de développement car les nuages de points des années suivantes présentent une structure aléatoire similaire.

Les résultats des tests graphiques montrent que les résidus présentent une forme aléatoire pour toutes les années de développement.

La statistique de test vaut 0,0077 et appartient à l'intervalle de confiance $[-0,1407; 0,1407]$. L'hypothèse H_0 de non-corrélation des résidus sur le triangle dans son ensemble par années de survenance est donc acceptée.

Les résultats du test graphique et du test statistique du ρ de Spearman permettent de valider l'hypothèse de variance pour la LoB income protection.

Les hypothèses (H2) et (H3) étant acceptée, le rejet de l'hypothèse (H1) nous amène donc à étudier la dépendance existante entre les années de survenance afin de pouvoir appliquer le modèle sans tenir compte du rejet de cette hypothèse. Le rejet de cette hypothèse est dû à la présence d'effets calendaires au sein des triangles de données. Un facteur calendaire pouvant être à l'origine de ces effets est l'inflation à laquelle nous allons nous intéresser dans la partie qui suit.

C. Triangles mis en « as-if »

Pour la LoB medical expenses, l'indice pris en compte est l'indice concernant les frais de santé : l'ONDAM (Objectif National de Dépenses d'Assurance Maladie). Pour la LoB income protection, l'indice pris en compte est l'indice d'inflation : le CPI (Consumer Price Index). Les triangles étudiés possèdent 26 années de survenance : de 1996 à 2021. L'année de référence est l'année 2021. Par conséquent, les triangles seront « as-if » 2021.

CY	CPI	ONDAM
1997	1,2%	1,5%
1998	0,7%	4,0%
1999	0,5%	2,6%
2000	1,7%	5,6%
2001	1,6%	5,6%
2002	1,9%	7,1%
2003	2,1%	6,4%
2004	2,1%	4,9%
2005	1,7%	4,0%
2006	1,7%	3,2%
2007	1,5%	4,0%
2008	2,8%	3,5%
2009	0,1%	3,6%
2010	1,5%	2,6%
2011	2,1%	2,7%
2012	2,0%	2,3%
2013	0,9%	2,2%
2014	0,5%	2,4%
2015	0,0%	2,1%
2016	0,2%	1,8%
2017	1,0%	2,2%
2018	1,9%	2,2%
2019	1,1%	2,6%
2020	0,5%	2,5%
2021	1,6%	4,60%

Figure 23 – Indices du CPI et de l'ONDAM

CY	CPI_taux	ONDAM_taux
1996	1,39	2,33
1997	1,37	2,29
1998	1,36	2,21
1999	1,35	2,15
2000	1,33	2,04
2001	1,31	1,93
2002	1,29	1,80
2003	1,26	1,69
2004	1,23	1,61
2005	1,21	1,55
2006	1,19	1,50
2007	1,17	1,44
2008	1,14	1,40
2009	1,14	1,35
2010	1,12	1,31
2011	1,10	1,28
2012	1,08	1,25
2013	1,07	1,22
2014	1,06	1,19
2015	1,06	1,17
2016	1,06	1,15
2017	1,05	1,12
2018	1,03	1,10
2019	1,02	1,07
2020	1,02	1,05
2021	1,00	1,00

Figure 22 – Taux d'actualisation en base 2021

Les taux du CPI et de l'ONDAM de 1997 à 2020 (figure 52) sont issus de l'INSEE. Pour l'année calendaire 2021, les taux sont retenus sont les taux estimés par le modèle d'inflation utilisé au sein du modèle interne partiel de Groupama.

Les taux d'actualisation (figure 53) utilisés pour avoir une vision 2021 des montants du triangle sont calculés à partir des indices de l'INSEE. Le processus de mise en as-if a été réalisé à partir d'un code R. Les triangles obtenus suite à la mise en as-if 2021 sont présentés en annexes.

Le test d'hypothèse entre années de survenance est effectué sur ces nouvelles données. Les résultats sont présentés ci-dessous.

LoB	Z	E(Z)	Var(Z)	Intervalle de confiance		Hypothèse
med.exp	34	121,6	28,2	111,2	132,0	Rejetée
inc.pr	41	122,4	28,4	112,0	132,9	Rejetée

Figure 24 – Résultats du test d'hypothèse d'indépendance entre années de survenance sur les triangles « as-if » 2021

Pour les deux LoB étudiées, le test d'absence d'effet calendaire est rejeté :

- Pour la la LoB medical expenses, la statistique de test vaut 34, ce qui n'appartient pas à l'intervalle de confiance [111,2 ; 132,0].
- Pour la LoB income protection, la statistique de test vaut 41, ce qui n'appartient pas à l'intervalle de confiance [112,0 ; 132,9].

Les résultats de ce test mènent à conclure que l'inflation n'est pas la cause du rejet de l'hypothèse d'indépendance entre années de survenance. Cette hypothèse étant rejetée, il est inutile de vérifier les autres hypothèses du modèle de Mack Chain-Ladder, à savoir, l'hypothèse de linéarité (H2) et l'hypothèse de variance (H3).

Le retraitement de l'inflation dans le triangle des paiements ne supprime pas les effets calendaires à l'origine du rejet du test d'indépendance. Nous allons alors étudier la dépendance présente au sein des triangles de données dans la partie suivante.

D. Ajout de dépendance entre années de survenance

En première approche, nous conservons le cadre de la modélisation de la méthode de Merz-Wuthrich répliquée par Bootstrap chez Groupama (Bootstrap à un an) et nous cherchons à étudier l'impact de l'ajout de dépendance entre années de survenance.

1. Choix de la copule

Lors de la partie théorique, plusieurs copules ont été présentées. Pour la réalisation de notre étude, notre choix s'est porté sur l'utilisation d'une copule gaussienne pour diverses raisons.

Tout d'abord, les copules Archimédiennes sont utilisées dans le cas des risques extrêmes, c'est-à-dire des risques coûteux pour l'entreprise en matière de SCR (catastrophes naturelles par exemple). En effet, cette famille de copule permet de capter une dépendance au niveau des queues de distribution. Dans notre cas, le risque de réserves est un risque de boni-mali autour de la moyenne des CDR et ne représente pas un risque extrême entraînant une perte considérable pour l'entreprise.

La méthode du Bootstrap à un an fait l'hypothèse que les résidus de Pearson sont de loi normale centrée réduite. De ce fait, la structure de dépendance recherchée est une structure normalisée. La copule gaussienne permet ainsi d'avoir une structure de dépendance centralisée ainsi que des résidus faibles et fortement corrélés entre eux.

2. Méthodes d'ajout de dépendance

Dans le cadre de notre étude, nous utilisons une copule gaussienne afin de simuler la diagonale N+1 lors du Bootstrap, dont les montants sont simulés à l'aide d'une loi log-normale de manière indépendante dans le modèle initial :

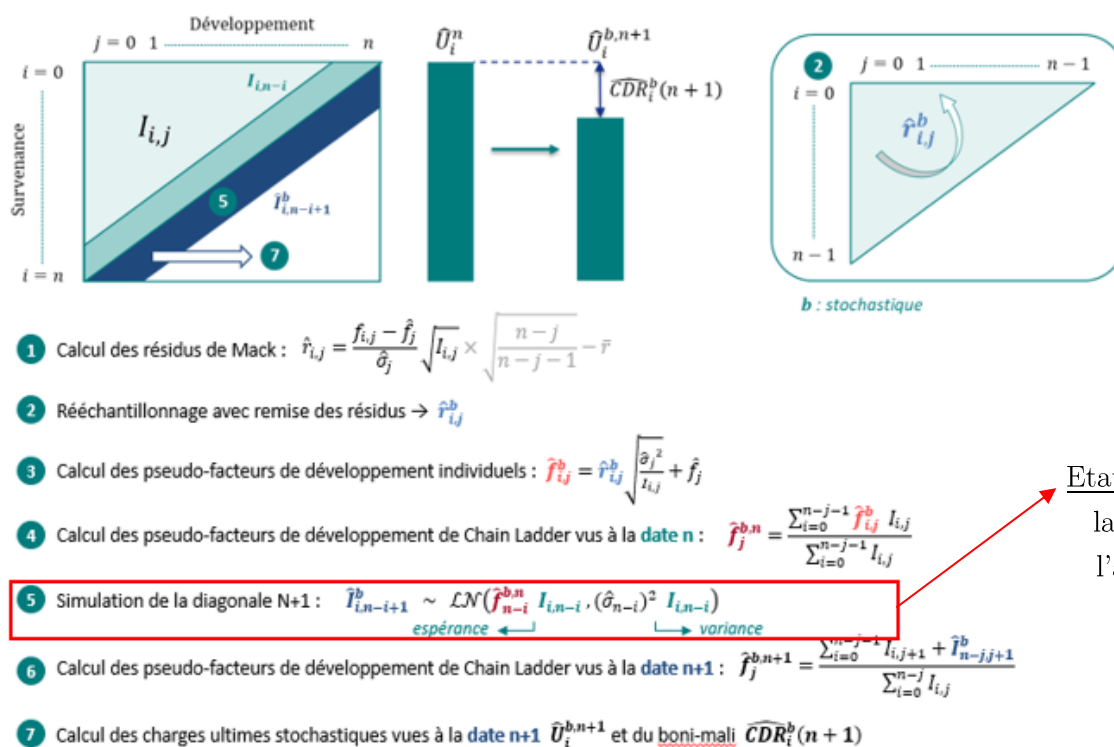


Figure 25 – Etape de la procédure Bootstrap (étape 5 modifiée)

Cette modification permet d'établir une dépendance a priori, c'est-à-dire que les montants sont corrélés avant projection à l'ultime et donc avant calcul de la volatilité à un an.

L'ajout de la copule gaussienne ne pouvant pas être testé sur le logiciel de provisionnement utilisé par Groupama, ce test a été réalisé à l'aide du logiciel R en reproduisant la méthode du Bootstrap à un an avec la modification de l'étape 5.

Nous comparons cette variante au scénario central, c'est-à-dire la méthode du Bootstrap à un an décrite en première partie, et également à une méthode d'impact a posteriori. L'impact a posteriori correspond à l'ajout de corrélation après calcul des volatilités à un an. La volatilité à un an totale est calculée par agrégation des volatilités individuelles à l'aide d'une matrice d'équi-corrélation. Une matrice d'équi-corrélation est une matrice de Toeplitz de la forme suivante :

$$\begin{pmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & 1 & \rho \\ \rho & \dots & \rho & 1 \end{pmatrix}$$

Où ρ est le coefficient de corrélation supposé unique et identique entre toutes les années de survenance. Les valeurs de corrélation testées sont identiques dans les deux cas de figure : 25%, 50%, 75%, 100% et ont été choisies de manière arbitraire.

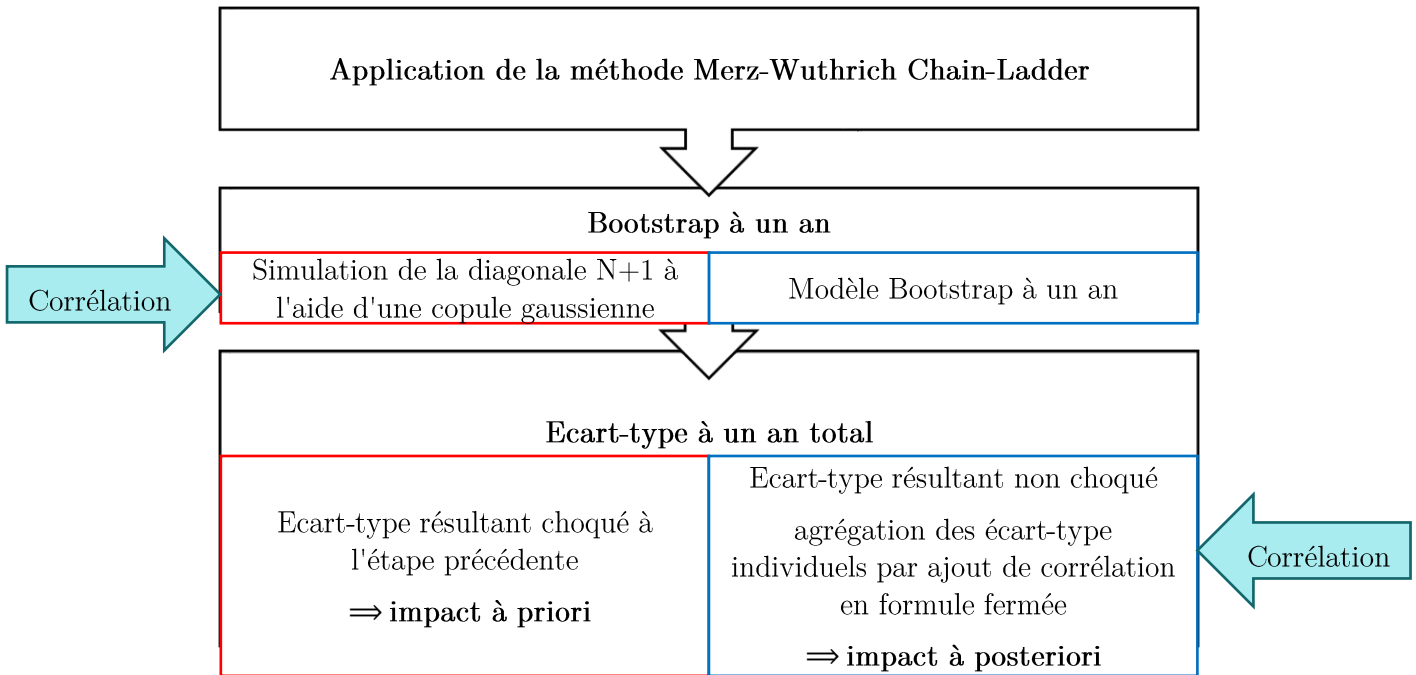


Figure 26 – Description des impacts de corrélation étudiés

Les termes à priori et à posteriori font appel à l'étape durant laquelle le choc de corrélation est appliqué :

- A priori : avant l'estimation de l'écart-type total à un an
- A posteriori : après l'estimation de l'écart-type total à un an

Impact 25%	MSEP central	MSEP à posteriori	MSEP à priori	% à posteriori	% à priori
med.exp.	14,0	14,4	8,7	2,7%	-38,1%
inc.pr.	21,5	26,8	31,5	24,4%	46,3%

Impact 50%	MSEP central	MSEP à posteriori	MSEP à priori	% à posteriori	% à priori
med.exp.	14,0	14,8	9,1	5,4%	-35,3%
inc.pr.	21,5	31,7	39,3	47,4%	82,6%

Impact 75%	MSEP central	MSEP à posteriori	MSEP à priori	% à posteriori	% à priori
med.exp.	14,0	15,1	9,5	8,0%	-32,5%
inc.pr.	21,5	36,0	45,9	67,3%	113,1%

Impact 100%	MSEP central	MSEP à posteriori	MSEP à priori	% à posteriori	% à priori
med.exp.	14,0	15,5	9,8	10,6%	-29,9%
inc.pr.	21,5	39,8	51,7	85,0%	140,4%

Figure 27 – Comparaison des MSEP à un an (en M€)

Les valeurs décrites dans le tableau ci-dessus représentent :

- MSEP Centrale : écart-type à un an total obtenu par application de la méthode du Bootstrap à un an.
- MSEP à posteriori : écart-type à un an total obtenu par agrégation des écart-types individuels à l'aide des matrices d'équi-corrélation.
- MSEP à priori : écart-type à un an total obtenu par simulation de la diagonale N+1 à l'aide d'une copule gaussienne.
- Les pourcentages correspondent aux pourcentages de hausse/baisse par rapport à la MSEP Centrale issue de la méthode du Bootstrap à un an.

Nous observons deux phénomènes assez différents entre les lignes d'activité étudiées :

Impact à priori :

Dans le cas de la LoB medical expenses, la simulation de la diagonale N+1 par une copule gaussienne entraîne une baisse de volatilité : -5,3M€ (-38,1%) (valeur de corrélation : 25%).

Dans le cas de la LoB income protection, cet impact entraîne une hausse de volatilité : +10M€ (+46,3%) (valeur de corrélation : 25%).

Les résultats obtenus semblent dépendre de la nature de la ligne d'activité. Des études complémentaires devront être menées afin d'expliquer la cause de ce phénomène et de savoir si les résultats reflètent la réalité.

Impact à posteriori :

Les résultats obtenus pour l'impact à posteriori sont cohérents du fait que les corrélations implicites du modèle pour les deux LoB sont inférieures aux valeurs de corrélations ajoutées. En effet, plus la corrélation entre les écart-type à un an individuels est forte plus l'écart-type à un an total sera dans ce cas élevé.

La corrélation implicite du modèle est obtenue par résolution d'un problème matriciel. Nous supposons que la volatilité totale à un an est obtenue par agrégation des volatilités individuelles par années de survenance et qu'il existe un unique coefficient de corrélation, noté ρ , entre toutes les années de survenance :

$$mse_{tot} = \sqrt{\sum_{i=1}^N mse_i^2 + 2 \sum_{1 \leq i < j \leq N} \rho \times mse_i \times mse_j}$$

Les volatilités à un an étant connues, nous construisons ainsi la matrice d'équi-corrélation.

Nous nous intéressons alors à la corrélation implicite entre les volatilités à un an lors de l'application de l'impact à priori. Afin d'estimer la corrélation implicite dans les deux cas de figure présentés précédemment, nous utilisons la méthode du GRG (Generalized Reduced Gradient) non linéaire du solveur Excel.

En notant :

- X : vecteur des écart-types individuels à un an de taille $N \times 1$
- Σ : matrice de corrélation de taille $N \times N$ et de la forme :

$$\begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & 1 & \rho \\ \rho & \cdots & \rho & 1 \end{pmatrix}$$

Où ρ : coefficient de corrélation inconnu à estimer

- mse_{tot} : l'écart-type à un an total estimé

Le solveur Excel estime une valeur de ρ telle que l'égalité suivante soit vraie :

$$mse_{tot} = (X^t \Sigma X)^{\frac{1}{2}}$$

Ainsi, le coefficient ρ estimé représente la corrélation implicite du modèle.

Nous nous intéressons alors à la corrélation implicite des deux modèles, à savoir le modèle du Bootstrap à un an appliqué et le modèle d'impact à priori par utilisation de la copule gaussienne pour la simulation de la diagonale N+1.

Corrélation Lob	Central	A priori	25% - sortie
med.exp.	0,2%	24,9%	-0,1%
inc.pr	3,1%	38,7%	13,7%

Corrélation Lob	Central	A priori	50% - sortie
med.exp.	0,2%	48,9%	-1,1%
inc.pr	3,1%	63,3%	13,3%

Corrélation Lob	Central	A priori	75% - sortie
med.exp.	0,2%	72,3%	-2,7%
inc.pr	3,1%	81,6%	6,6%

Corrélation Lob	Central	A priori	100% - sortie
med.exp.	0,2%	94,9%	-5,1%
inc.pr	3,1%	95,9%	-4,1%

Figure 28 – Comparaison des corrélations implicites entre MSEP

Les valeurs décrites dans le tableau ci-dessus représentent :

- Central : corrélation implicite entre les écart-type à un an individuels issus du scénario central, qui est pour rappel, la méthode du Bootstrap à un an utilisée dans le modèle interne partiel.
- A priori : corrélation implicite entre les écart-type individuels à un an issus de l'impact à priori (simulation de la diagonale N+1 à l'aide d'une copule gaussienne).
- Valeur corrélation% - sortie : différence entre le paramètre de corrélation ajouté dans la copule (25%, 50%, 75% et 100%) et la corrélation implicite calculée entre les écart-type individuels à un an issus de l'impact à priori.

Les corrélations implicites du scénario central sont bien inférieures aux valeurs de corrélations ajoutées, comme mentionné dans les commentaires de la figure 15.

En théorie, sous l'hypothèse d'indépendance, les corrélations en sortie de la méthode d'impact à priori devraient être égales aux valeurs testées, à savoir : 25%, 50%, 75% et 100%. Ce résultat est vérifié pour la LoB medical expenses mais pas pour la LoB income protection. En effet, nous pouvons observer que pour la LoB medical expenses, la corrélation implicite entre les écart-type à un an est égale à la valeur de corrélation testée (par exemple : 25% en entrée → 24,9% en sortie). Ce résultat n'est pas vérifié au

niveau de la LoB income protection, où nous observons une hausse de corrélation implicite par rapport à la corrélation d'entrée (par exemple : 25% en entrée → 38,7% en sortie).



Figure 29 – Comportement de la corrélation lors l'impact à priori

Lors de la projection à l'ultime, la corrélation est amplifiée. Les probables explications de cet effet sont les suivantes :

- Spécificité de la ligne d'activité : pour les valeurs 25% et 50%, l'augmentation de corrélation est du même ordre de grandeur à savoir, respectivement, 13,7% et 13,3% pour la LoB income protection. Les valeurs de corrélation testées plus élevés (75% et 100%) atténuent cette déformation.
- La longueur de la branche : lorsqu'une branche est longue, le choc de corrélation appliqué sur la diagonale N+1 est plus sévèrement propagé sur les corrélations des volatilités à un an.
- Les pseudo facteurs de développement utilisés lors de la projection à l'ultime :

4 Calcul des pseudo-facteurs de développement de Chain Ladder vus à la date n : $\hat{f}_j^{b,n} = \frac{\sum_{i=0}^{n-j-1} \hat{f}_{ij}^b I_{i,j}}{\sum_{i=0}^{n-j-1} I_{i,j}}$

5 Simulation de la diagonale N+1 : $\hat{I}_{i,n-i+1}^b \sim \mathcal{LN}(\underbrace{\hat{f}_{n-i}^{b,n}}_{\text{espérance}} I_{i,n-i}, \underbrace{(\hat{\sigma}_{n-i})^2}_{\text{variance}} I_{i,n-i})$

6 Calcul des pseudo-facteurs de développement de Chain Ladder vus à la date n+1 : $\hat{f}_j^{b,n+1} = \frac{\sum_{i=0}^{n-j-1} I_{i,j+1} + \hat{I}_{n-jj+1}^b}{\sum_{i=0}^{n-j} I_{i,j}}$

Figure 30 – Etapes 4 à 6 de la méthode du Bootstrap à un an

Lors de la procédure Bootstrap, de nouveaux coefficients de développement sont calculés en tenant compte de la nouvelle information présente dans la diagonale N+1 (étape 6 ci-dessus). Ces coefficients intègrent donc la dépendance créée par l'utilisation de la copule lors de l'étape 5 de la méthode du Bootstrap et sont utilisés pour la projection à l'ultime de la diagonale N+1.

Par conséquent, le choc de corrélation peut être comptabilisé deux fois. Nous avons utilisé les coefficients de développement calculés avant simulation de la diagonale N+1 (**étape 4 ci-dessus**) afin de comparer la corrélation implicite entre les écart-type en fonction des coefficients de développement utilisés lors de l'impact à priori.

Corrélation Lob	Central	DF 6	DF 4
med.exp.	0,2%	24,9%	23,7%
inc.pr	3,1%	38,7%	26,2%

Figure 31 – Comparaison des corrélations implicites (25%)

Les valeurs décrites dans le tableau ci-dessus représentent :

- Centrale : corrélation implicite entre les écart-type à un an individuels issus du scénario central, qui est pour rappel, la méthode du Bootstrap à un an utilisée dans le modèle interne partiel.
- DF 6 (Development Factor 6) : corrélation implicite entre les écart-type individuels à un an issus de l'impact à priori (simulation de la diagonale N+1 à l'aide d'une copule gaussienne) en utilisant les pseudo-coefficients de développement calculés lors de l'**étape 6** de la procédure Bootstrap (tenant compte de l'information de la diagonale N+1).
- DF 4 (Development Factor 4) : corrélation implicite entre les écart-type individuels à un an issus de l'impact à priori (simulation de la diagonale N+1 à l'aide d'une copule gaussienne) en utilisant les pseudo-coefficients de développement calculés lors de l'**étape 4** de la procédure Bootstrap (ne tenant pas compte de l'information de la diagonale N+1).

Pour la LoB income protection, l'utilisation des pseudo-coefficients de Chain-Ladder avant simulation de la diagonale N+1 (étape 4) diminue la corrélation en sortie de 12,5% et permet de s'approcher de 25% (38,7% vs 26,2%).

Impact 25% Lob	MSEP central	MSEP DF 6	MSEP DF 4	% DF 6	% DF 4
med.exp.	14,0	8,7	8,6	-38,1%	-38,4%
inc.pr.	21,5	31,5	27,7	46,3%	28,8%

Figure 32 – Comparaison des volatilités à un an en fonction des pseudo-facteurs de développement utilisés

Les valeurs décrites dans le tableau ci-dessus représentent :

- MSEP Centrale : écart-type à un an total obtenu par application de la méthode du Bootstrap à un an.
- MSEP DF 6 : écart-type à un an total issu de l'impact à priori (simulation de la diagonale N+1 à l'aide d'une copule gaussienne) en utilisant les pseudo-coefficients de développement calculés

lors de l'**étape 6** de la procédure Bootstrap (tenant compte de l'information de la diagonale N+1).

- MSEP DF 4 : écart-type à un an total issus de l'impact à priori (simulation de la diagonale N+1 à l'aide d'une copule gaussienne) en utilisant les pseudo-coefficients de développement calculés lors de l'**étape 4** de la procédure Bootstrap (ne tenant pas compte de l'information de la diagonale N+1).
- Les pourcentages correspondent aux pourcentages de hausse/baisse de volatilité par rapport à la MSEP Central issue la méthode du Bootstrap à un an.

La volatilité à un an de la LoB income protection diminue de 3,8M€ dans le cas où les pseudo-coefficients n'intègrent pas l'information de la diagonale N+1 (**étape 4**) par rapport au cas où les pseudo-coefficients sont calculés après simulation de la diagonale N+1 (**étape 6**) (27,7M€ vs 31,5M€).

L'ajout de dépendance dans le modèle utilisé à savoir, la méthode du Bootstrap à un an, via une copule gaussienne ne permet pas d'obtenir des résultats exploitables sans études supplémentaires.

L'hypothèse de l'unicité du coefficient de corrélation est assez forte. En effet, deux années de survenance séparées de 10 ans par exemple ont peu de chance d'être corrélées de manière identique à deux années de survenance voisines ou séparées de 1 an ou 2 ans.

Nous allons nous intéresser au coefficient de corrélation des rangs de Spearman afin d'étudier la dépendance entre chaque couple d'années de survenance.

3. Corrélations entre années de survenance

Dans le cas étudié, la corrélation entre années de survenance est estimée à l'aide du ρ de Spearman. Les matrices de corrélations des rangs présentées ci-dessous représentent les corrélations entre les boni-mali obtenus après application de la méthode du Bootstrap à un an et sont obtenues à l'aide de l'outil de provisionnement ResQ.

Les boni-mali sont supposés nuls durant la première année de survenance. Par conséquent, l'année de survenance 1996 n'apparaît pas dans les matrices de corrélation des rangs.

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
2	100%	50%	49%	50%	52%	53%	54%	1%	0%	2%	1%	2%	0%	0%	1%	-1%	-1%	0%	1%	2%	-1%	-1%	0%	0%	0%
3	50%	100%	58%	59%	60%	62%	63%	1%	0%	2%	-1%	0%	-1%	0%	0%	-1%	-1%	0%	0%	0%	0%	-1%	0%	1%	0%
4	49%	58%	100%	60%	62%	64%	65%	1%	1%	2%	0%	1%	0%	1%	1%	0%	-1%	1%	0%	1%	0%	-1%	-1%	0%	0%
5	50%	59%	60%	100%	62%	64%	65%	1%	0%	3%	1%	1%	1%	0%	0%	-1%	-2%	0%	0%	1%	0%	-1%	0%	0%	-1%
6	52%	60%	62%	62%	100%	65%	67%	1%	0%	3%	0%	1%	1%	1%	1%	0%	-1%	1%	1%	1%	0%	0%	-1%	-1%	-1%
7	53%	62%	64%	64%	65%	100%	68%	2%	1%	3%	0%	1%	-1%	0%	0%	-1%	-1%	0%	0%	1%	0%	-1%	0%	0%	-1%
8	54%	63%	65%	65%	67%	68%	100%	1%	0%	3%	0%	1%	0%	0%	0%	0%	-1%	0%	1%	1%	0%	0%	0%	0%	0%
9	1%	1%	1%	1%	1%	2%	1%	100%	8%	29%	1%	11%	6%	3%	1%	1%	-1%	-1%	1%	0%	0%	0%	-1%	-1%	-1%
10	0%	0%	1%	0%	0%	1%	0%	8%	100%	44%	2%	16%	9%	4%	2%	3%	1%	0%	1%	1%	0%	0%	1%	-1%	0%
11	2%	2%	2%	3%	3%	3%	3%	29%	44%	100%	2%	13%	7%	4%	2%	3%	1%	0%	1%	0%	0%	1%	1%	-1%	0%
12	1%	-1%	0%	1%	0%	0%	0%	1%	2%	2%	100%	71%	40%	20%	15%	9%	5%	2%	3%	2%	1%	0%	-1%	0%	0%
13	2%	0%	1%	1%	1%	1%	1%	11%	16%	13%	71%	100%	33%	17%	12%	8%	5%	1%	3%	2%	0%	0%	-1%	0%	-1%
14	0%	-1%	0%	1%	1%	-1%	0%	6%	9%	7%	40%	33%	100%	12%	10%	7%	4%	1%	2%	1%	1%	-1%	-1%	0%	0%
15	0%	0%	1%	0%	1%	0%	0%	3%	4%	4%	20%	17%	12%	100%	11%	7%	3%	1%	3%	1%	0%	0%	0%	-1%	-2%
16	1%	0%	1%	0%	1%	0%	0%	1%	2%	2%	15%	12%	10%	11%	100%	5%	3%	2%	2%	3%	-1%	-1%	0%	-1%	0%
17	-1%	-1%	0%	-1%	0%	-1%	0%	1%	3%	3%	9%	8%	7%	7%	5%	100%	5%	2%	5%	3%	-1%	0%	1%	1%	1%
18	-1%	-1%	-1%	-2%	-1%	-1%	-1%	1%	1%	1%	5%	5%	4%	3%	3%	5%	100%	3%	6%	1%	1%	0%	-1%	0%	0%
19	0%	0%	1%	0%	1%	0%	0%	-1%	0%	1%	2%	1%	1%	1%	2%	2%	3%	1%	2%	2%	2%	3%	0%	1%	1%
20	1%	0%	0%	0%	1%	0%	1%	-1%	1%	1%	3%	3%	2%	3%	2%	5%	6%	26%	100%	10%	3%	2%	0%	0%	2%
21	2%	0%	1%	1%	1%	1%	1%	1%	1%	0%	2%	2%	1%	1%	3%	3%	4%	18%	10%	100%	1%	5%	0%	-2%	0%
22	-1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	1%	0%	-1%	1%	2%	3%	1%	2%	100%	5%	3%	0%	0%
23	-1%	-1%	-1%	-1%	0%	-1%	0%	0%	0%	0%	0%	0%	-1%	0%	-1%	0%	0%	3%	2%	2%	5%	100%	2%	0%	1%
24	0%	0%	-1%	0%	-1%	0%	0%	-1%	1%	1%	-1%	-1%	-1%	0%	0%	1%	-1%	0%	0%	0%	3%	2%	100%	2%	-1%
25	0%	1%	0%	0%	-1%	0%	0%	-1%	-1%	-1%	0%	0%	0%	-1%	-1%	1%	0%	1%	0%	-2%	0%	0%	2%	100%	0%
26	0%	0%	0%	-1%	-1%	-1%	0%	-1%	0%	0%	0%	0%	-1%	-2%	0%	1%	0%	1%	2%	0%	0%	1%	-1%	0%	100%

Figure 33 – Matrice de corrélations des rangs entre boni-mali (medical expenses)

Pour la LoB medical expenses, quatre sous matrices de corrélations se distinguent (encadrées en bleu ci-dessus). Certaines années de survénance sont non corrélées ou très faiblement, cela permet donc de cibler les différentes années de survénance à l'origine du rejet de l'hypothèse d'indépendance.

Les chocs de corrélation pourront ainsi être plus adaptés et tendre vers la réalité. En effet, les différents blocs ne présentent pas les mêmes degrés de corrélation : les huit premières années de survénance semblent être plus corrélées que les années de survénance 9, 10 et 11. De ce fait, la méthode de l'impact à posteriori présentée dans la partie précédente est trop prudente par rapport à la réalité.

Il est possible d'adapter cet impact à posteriori en construisant une matrice de corrélation par blocs. Chaque bloc serait une matrice d'équi-corrélation avec un coefficient ρ adapté au degré de corrélation existant entre les années de survénance. La matrice de corrélation serait alors de la forme :

$$\begin{pmatrix} A & 0 & 0 & 0 \\ 0 & B & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & D \end{pmatrix}$$

Où A, B, C et D sont les matrices d'équi-corrélation construites à partir de l'identification des blocs (en bleu) présents dans la matrice de corrélation du ρ de Spearman. Chacune de ces matrices possèdera alors un coefficient de corrélation unique et adapté au degré de liaison entre les années de survénance.

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
2	100%	26%	43%	29%	38%	41%	35%	33%	40%	28%	29%	27%	29%	17%	20%	11%	15%	15%	14%	10%	7%	7%	5%	2%	2%
3	26%	100%	64%	43%	56%	61%	52%	48%	59%	41%	43%	41%	43%	23%	28%	16%	21%	22%	19%	13%	11%	8%	6%	3%	2%
4	43%	64%	100%	42%	54%	59%	50%	47%	57%	39%	41%	39%	42%	23%	27%	16%	21%	22%	18%	14%	10%	8%	6%	2%	1%
5	29%	43%	42%	100%	52%	56%	48%	45%	54%	37%	39%	37%	40%	22%	26%	15%	21%	20%	17%	12%	9%	8%	5%	2%	1%
6	38%	56%	54%	52%	100%	60%	52%	47%	58%	41%	43%	41%	43%	24%	28%	17%	21%	22%	19%	14%	11%	9%	6%	3%	1%
7	41%	61%	59%	56%	60%	100%	53%	49%	61%	41%	43%	41%	45%	24%	29%	17%	22%	23%	19%	14%	11%	9%	7%	3%	2%
8	35%	52%	50%	48%	52%	53%	100%	45%	54%	39%	39%	37%	41%	22%	26%	16%	20%	21%	18%	13%	10%	8%	6%	2%	2%
9	33%	48%	47%	45%	47%	49%	45%	100%	52%	37%	38%	37%	39%	21%	25%	15%	20%	19%	17%	12%	9%	8%	4%	2%	1%
10	40%	59%	57%	54%	58%	61%	54%	52%	100%	41%	43%	42%	44%	25%	29%	17%	22%	23%	19%	14%	10%	9%	7%	2%	2%
11	28%	41%	39%	37%	41%	41%	39%	37%	41%	100%	35%	33%	35%	21%	24%	14%	18%	19%	16%	11%	8%	7%	5%	3%	1%
12	29%	43%	41%	39%	43%	43%	39%	38%	43%	35%	100%	34%	37%	21%	24%	15%	19%	19%	16%	11%	9%	7%	4%	3%	1%
13	27%	41%	39%	37%	41%	41%	37%	37%	42%	33%	34%	100%	36%	20%	24%	13%	19%	18%	16%	11%	8%	7%	4%	2%	1%
14	29%	43%	42%	40%	43%	45%	41%	39%	44%	35%	37%	36%	100%	20%	24%	15%	20%	19%	16%	11%	10%	7%	5%	3%	1%
15	17%	23%	23%	22%	24%	24%	22%	21%	25%	21%	21%	20%	20%	100%	20%	12%	15%	15%	14%	9%	7%	6%	4%	1%	0%
16	20%	28%	27%	26%	28%	29%	26%	25%	29%	24%	24%	24%	24%	20%	100%	11%	15%	16%	14%	10%	7%	5%	5%	2%	1%
17	11%	16%	16%	15%	17%	17%	16%	15%	17%	14%	15%	13%	15%	12%	11%	100%	16%	16%	15%	10%	5%	7%	5%	3%	1%
18	15%	21%	21%	21%	21%	22%	20%	20%	22%	18%	19%	19%	20%	15%	15%	16%	100%	17%	13%	10%	7%	6%	4%	3%	1%
19	15%	22%	22%	20%	22%	23%	21%	19%	23%	19%	18%	19%	15%	16%	16%	17%	100%	14%	10%	7%	7%	3%	2%	1%	
20	14%	19%	18%	17%	19%	19%	18%	17%	19%	16%	16%	16%	16%	14%	14%	15%	13%	14%	100%	10%	9%	6%	5%	2%	2%
21	10%	13%	14%	12%	14%	14%	13%	12%	14%	11%	11%	11%	11%	9%	10%	10%	10%	10%	10%	100%	7%	7%	4%	0%	0%
22	7%	11%	10%	9%	11%	11%	10%	9%	10%	8%	9%	8%	10%	7%	7%	5%	7%	9%	7%	7%	100%	8%	6%	3%	1%
23	7%	8%	8%	8%	9%	9%	8%	8%	9%	7%	7%	7%	7%	6%	5%	7%	6%	7%	6%	7%	8%	100%	6%	2%	2%
24	5%	6%	6%	5%	6%	7%	6%	4%	7%	5%	4%	5%	4%	4%	5%	4%	3%	5%	4%	6%	6%	100%	4%	0%	0%
25	2%	3%	2%	2%	3%	3%	2%	2%	2%	3%	3%	2%	3%	1%	2%	3%	3%	2%	2%	0%	3%	2%	4%	100%	2%
26	2%	2%	1%	1%	2%	2%	1%	2%	1%	1%	1%	1%	1%	0%	1%	1%	1%	1%	2%	0%	1%	2%	0%	2%	100%

Figure 34 – Matrice de corrélations des rangs entre boni-mali (income protection)

Pour la LoB income protection, toutes les années de survénance semblent être corrélées entre elles à l'exception des trois dernières qui le sont peu ou pas du tout. Une sous-matrice se distingue (encadrée en noir ci-dessus) faisant intervenir des degrés de corrélation différents (encadrés en bleu et en jaune).

La méthode de l'impact à posteriori présentée dans la partie précédente est trop violente par rapport à la réalité. En effet, les années de survénance ne sont pas corrélées à l'aide d'un unique coefficient.

Il est possible d'adapter cet impact à posteriori en construisant une matrice de corrélation donc les coefficients de corrélation sont fixés tels qu'ils représentent au mieux le degré de liaison entre les différentes années de survénance. Par exemple, fixer une valeur de corrélation de la deuxième à la quatorzième année de survénance (encadrées en bleu) puis une seconde valeur plus faible de la quinzième à la vingt-deuxième année de survénance (encadrées en jaune). La matrice de corrélation serait alors de la forme :

$$\begin{pmatrix} 1 & \rho & \rho & \nu & \nu & 0 \\ \rho & 1 & \rho & \nu & \nu & 0 \\ \rho & \rho & 1 & \nu & \nu & 0 \\ \nu & \nu & \nu & 1 & \nu & 0 \\ \nu & \nu & \nu & \nu & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Où ρ et ν sont les corrélations retenues tels que $\rho > \nu$.

Le découpage en deux niveaux de corrélation (bleu et jaune) ne repose sur aucun test et n'est pas justifié. Par conséquent, il est possible de procéder à un découpage plus fin pour être plus précis dans le choix des valeurs de corrélation.

4. Limites

Les méthodes présentées autour de l'ajout de dépendance entre les années de survenance à travers les écart-type individuels à un an présentent des limites.

- Les choix des différentes valeurs de corrélation : Les différentes valeurs testées ont été fixées de manière arbitraire et leur fiabilité n'est pas prouvée. Par conséquent, il n'est pas possible de faire un choix justifié entre les différentes valeurs testées mais uniquement d'évaluer l'impact sur l'écart-type à un an.
- Méthodes instables : Ces méthodes ont des impacts différents en fonction de la ligne d'activité étudié. Cette différence pourrait être liée à la spécificité de chaque ligne d'activité mais aussi à l'instabilité de la méthode et nécessiterait des études complémentaires.

Les résultats obtenus par application d'une copule gaussienne ne permettent pas de retenir cette méthode. Cependant, la méthode d'agrégation via une matrice d'équi-corrélation des volatilités par années de survenance nous mène vers une potentielle solution : les matrices de corrélations par blocs. Cette variante pourrait répondre à notre problématique à condition d'approfondir l'étude du comportement des années de survenance entre elles (points avec les gestionnaires sinistres, backtesting, ...) conduisant à des choix pertinents de niveau de corrélation.

Après avoir étudié l'impact de différentes méthodes d'ajout de corrélation, il est intéressant de tester une méthode substituant à la méthode Chain-Ladder. Nous avons fait le choix d'étudier l'application des modèles linéaires généralisés à travers la famille des lois Tweedie.

E. Modèles linéaires généralisés

1. Résultats

L'estimation des paramètres est réalisée à l'aide de la fonction `tweedieReserve()` incluse dans le package `ChainLadder` du logiciel R.

Les résultats suivants correspondent à l'application des deux modèles GLM étudiés qui sont, pour rappel, les GLM Poisson et Tweedie que l'on compare aux résultats issus du modèle de base appliqué qui est la

méthode du Bootstrap à un an. Les GLM étant appliqués sur les triangles des paiements incrémentaux, nous utiliserons le triangle des paiements incrémentaux de la LoB income protection pour l'application du GLM ainsi que la méthode du Bootstrap.

Les résultats présentés ci-dessous correspondent à l'estimation des paramètres du modèle GLM Poisson des deux LoB étudiées.

Paramètres	Estimation	Paramètres	Estimation
(Intercept)	19,7507	(Intercept)	19,7507
(origin)2	0,0180	(dev)1	-1,4310
(origin)3	0,0574	(dev)2	-5,7188
(origin)4	0,0945	(dev)3	-7,9121
(origin)5	0,1484	(dev)4	-9,1649
(origin)6	0,1850	(dev)5	-9,8804
(origin)7	0,2281	(dev)6	-11,2802
(origin)8	0,2875	(dev)7	-11,6981
(origin)9	0,3375	(dev)8	-10,7276
(origin)10	0,3736	(dev)9	-13,4303
(origin)11	0,4074	(dev)10	-13,0744
(origin)12	0,4525	(dev)11	-13,6872
(origin)13	0,5650	(dev)12	-14,0613
(origin)14	0,6233	(dev)13	-14,6652
(origin)15	0,6900	(dev)14	-16,0674
(origin)16	0,7655	(dev)15	-13,4798
(origin)17	0,7678	(dev)16	-16,4473
(origin)18	0,7548	(dev)17	-15,0116
(origin)19	0,7433	(dev)18	-15,7297
(origin)20	0,7300	(dev)19	-19,5017
(origin)21	0,7093	(dev)20	-29,1372
(origin)22	0,7089	(dev)21	-29,1170
(origin)23	0,7432	(dev)22	-29,0958
(origin)24	0,7684	(dev)23	-29,0785
(origin)25	0,7620	(dev)24	-29,0623
(origin)26	0,8791	(dev)25	-29,0533

Figure 35 – Estimation des paramètres du GLM Poisson (medical expenses)

Paramètres	Estimation
(Intercept)	17,7538
(origin)2	0,0343
(origin)3	0,0894
(origin)4	0,0842
(origin)5	0,1431
(origin)6	0,1872
(origin)7	0,1192
(origin)8	0,0525
(origin)9	0,0621
(origin)10	0,1173
(origin)11	0,1437
(origin)12	0,1324
(origin)13	0,1611
(origin)14	0,2274
(origin)15	0,2883
(origin)16	0,3088
(origin)17	0,3422
(origin)18	0,3731
(origin)19	0,3761
(origin)20	0,4020
(origin)21	0,4269
(origin)22	0,5345
(origin)23	0,5662
(origin)24	0,6336
(origin)25	0,5853
(origin)26	0,5452

Paramètres	Estimation
(Intercept)	17,7538
(dev)1	0,2430
(dev)2	-0,7973
(dev)3	-1,3534
(dev)4	-2,1222
(dev)5	-2,8988
(dev)6	-3,4204
(dev)7	-4,0636
(dev)8	-4,2403
(dev)9	-4,8453
(dev)10	-5,0134
(dev)11	-5,1764
(dev)12	-5,3266
(dev)13	-5,4278
(dev)14	-5,7669
(dev)15	-5,5432
(dev)16	-5,9338
(dev)17	-5,9308
(dev)18	-5,9298
(dev)19	-6,1014
(dev)20	-5,7295
(dev)21	-6,1437
(dev)22	-6,0544
(dev)23	-6,0646
(dev)24	-5,6836
(dev)25	-7,1285

Figure 36 – Estimation des paramètres du GLM Poisson (income protection)

Dans le cas du GLM Tweedie, le paramètre p de puissance de variance est estimé par maximum de vraisemblance.

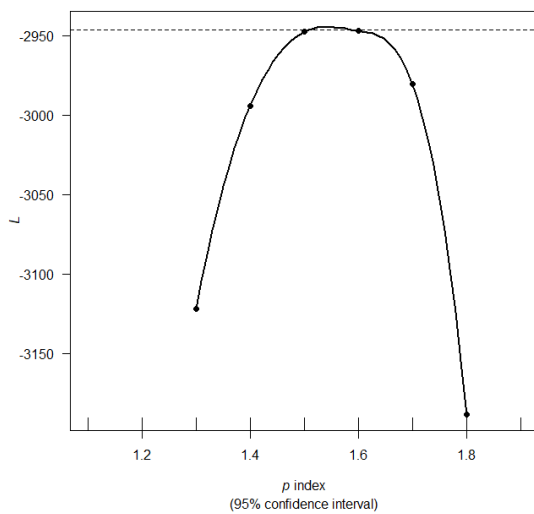


Figure 37 – Vraisemblance en fonction de la valeur du paramètre p (medical expenses)

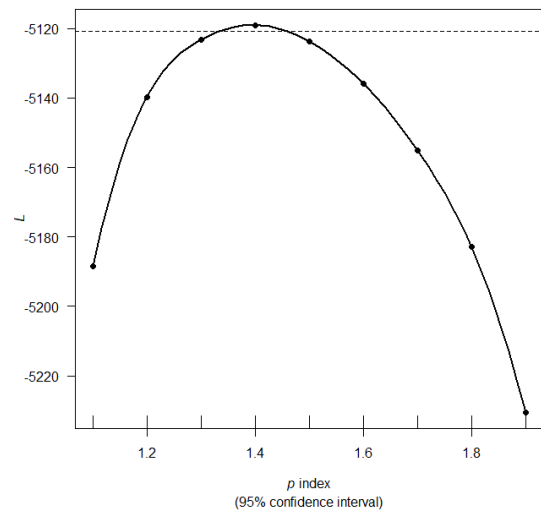


Figure 38 – Vraisemblance en fonction de la valeur du paramètre p (income protection)

LoB	p
med.exp.	1,5
inc.pr.	1,4

Figure 39 – Paramètres de puissance de variance

Le GLM Poisson correspond à un GLM Tweedie dont le paramètre de puissance p de variance vaut 1. Les paramètres de variance p estimés par maximum de vraisemblance pour les deux LoB étudiées sont supérieurs à 1. De ce fait, nous nous attendons à obtenir des résultats différents lors de l'application de ces deux GLM.

Après estimation de ce paramètre, la loi du modèle est définie et l'estimation des paramètres du modèle peut avoir lieu. Les résultats suivants correspondent à l'estimation des paramètres du modèle GLM Tweedie des deux LoB étudiées.

Paramètres	Estimation	Paramètres	Estimation
(Intercept)	19,7052	(Intercept)	19,7052
(origin)2	0,0220	(dev)1	-1,4317
(origin)3	0,0674	(dev)2	-5,7442
(origin)4	0,1178	(dev)3	-7,9464
(origin)5	0,2483	(dev)4	-9,2276
(origin)6	0,2286	(dev)5	-9,8493
(origin)7	0,2503	(dev)6	-11,3797
(origin)8	0,3687	(dev)7	-11,7938
(origin)9	0,3504	(dev)8	-10,7316
(origin)10	0,3924	(dev)9	-13,4290
(origin)11	0,4611	(dev)10	-13,0767
(origin)12	0,4660	(dev)11	-13,7596
(origin)13	0,5878	(dev)12	-14,0570
(origin)14	0,6637	(dev)13	-14,6418
(origin)15	0,7193	(dev)14	-15,9556
(origin)16	0,7983	(dev)15	-13,5195
(origin)17	0,7957	(dev)16	-16,3396
(origin)18	0,8104	(dev)17	-14,9267
(origin)19	0,8215	(dev)18	-15,6692
(origin)20	0,8185	(dev)19	-19,5417
(origin)21	0,7708	(dev)20	-43,1220
(origin)22	0,7702	(dev)21	-43,0992
(origin)23	0,8192	(dev)22	-43,0596
(origin)24	0,8305	(dev)23	-43,0377
(origin)25	0,8410	(dev)24	-43,0188
(origin)26	0,9246	(dev)25	-43,0078

Figure 40 – Estimation des paramètres du GLM Tweedie (medical expenses)

Paramètres	Estimation	Paramètres	Estimation
(Intercept)	17,6828	(Intercept)	17,6828
(origin)2	0,1701	(dev)1	0,2395
(origin)3	0,1491	(dev)2	-0,8165
(origin)4	0,1368	(dev)3	-1,3703
(origin)5	0,1992	(dev)4	-2,1334
(origin)6	0,2571	(dev)5	-2,9127
(origin)7	0,1242	(dev)6	-3,4276
(origin)8	0,0975	(dev)7	-4,0661
(origin)9	0,0795	(dev)8	-4,2373
(origin)10	0,1394	(dev)9	-4,8405
(origin)11	0,2337	(dev)10	-4,9994
(origin)12	0,1976	(dev)11	-5,1651
(origin)13	0,2325	(dev)12	-5,3155
(origin)14	0,3136	(dev)13	-5,4063
(origin)15	0,3608	(dev)14	-5,7499
(origin)16	0,4049	(dev)15	-5,5306
(origin)17	0,4466	(dev)16	-5,9096
(origin)18	0,4634	(dev)17	-5,9164
(origin)19	0,4972	(dev)18	-5,9183
(origin)20	0,4957	(dev)19	-6,0836
(origin)21	0,5521	(dev)20	-5,7171
(origin)22	0,6400	(dev)21	-6,1364
(origin)23	0,6809	(dev)22	-6,0578
(origin)24	0,7331	(dev)23	-6,0660
(origin)25	0,6544	(dev)24	-5,7000
(origin)26	0,6162	(dev)25	-7,0575

Figure 41 – Estimation des paramètres du GLM Tweedie (income protection)

Les paramètres $(origin)n$ correspondent aux effets des ligne l_i qui décrivent donc le comportement des années de de survenance.

Les paramètres $(dev)n$ correspondent aux effets des colonnes qui décrivent donc le comportement des années de développement.

Afin de comparer les trois méthodes entre elles, nous commençons par comparer les provisions Best Estimate (provisions BE) et les écart-type total à un an.

med.exp.	Central	GLM Poisson	GLM Tweedie
Provisions BE (k€)	225 071	225 201	224 974
MSEP (k€)	14 018	5 855	34 307

inc.pr.	Central	GLM Poisson	GLM Tweedie
Provisions BE (k€)	442 355	442 396	443 308
MSEP (k€)	17 350	15 080	25 084

Figure 42 – Provisions BE et écart-type à un an

Les valeurs décrites dans les tableaux ci-dessus représentent :

- Central : résultats issus du modèle utilisé à savoir la méthode du Bootstrap à un an servant de référence lors des analyses.
- GLM Poisson : résultats issus de l'application d'un GLM Poisson
- GLM Tweedie : résultats issus de l'application d'un GLM Tweedie

Les résultats attendus suite à l'application du GLM Poisson devraient être proches des résultats obtenus par la méthode utilisée actuellement au sein du modèle interne. En effet, la méthode Chain-Ladder et le GLM Poisson convergent vers le même résultat.

Les résultats sont considérés équivalents à la méthode du Bootstrap à un an utilisée au sein du modèle si l'écart représente $\pm 3\%$ des résultats centraux.

L'analyse des résultats obtenus pour la LoB medical expenses est effectuée ci-dessous.

GLM Poisson (med.exp.) :

- Les provisions Best Estimate sont proches du résultat de la méthode du Bootstrap à un an utilisée (+130 k€).
- L'écart-type à un an est quasiment divisé par trois (-8163 k€).

La différence significative entre l'écart-type Central et l'écart-type du GLM Poisson peut être expliquée par les données du triangle utilisé. Le triangle des paiements incrémentaux de cette ligne d'activité comporte un grand nombre de montants négatifs. La présence de valeurs négatives rend l'utilisation des GLM impossible. Pour pouvoir appliquer le GLM, les montants négatifs ont été remplacés par des valeurs nulles. Par conséquent, l'estimation des paramètres est de moins bonnes qualités, ce qui amène à une estimation du triangle inférieur et des paiements ultimes moins bonnes.

GLM Tweedie (med.exp.) :

- Les provisions Best Estimate sont proches des provisions Best Estimate Central (-227 k€).
- L'écart-type à un an est quasiment doublé (+20289 k€).

La différence significative entre l'écart-type Central et l'écart-type du GLM Tweedie peut être expliquée pour les mêmes raisons précédentes. La qualité du GLM est réduite en raison du nombre important de montants négatifs présents dans les données.

L'analyse des résultats obtenus pour la LoB income protection est effectuée à son tour.

GLM Poisson (inc.pr.) :

- Les provisions Best Estimate sont proches du résultat de la méthode du Bootstrap à un an utilisée (+44 k€).
- L'écart-type à un an diminue (-2270 k€).

La diminution de l'écart-type par application du GLM Poisson est moins forte que la baisse observée pour la LoB medical expenses, cependant elle n'est pas négligeable. La LoB income protection comporte moins de montants négatifs que la LoB medical expenses, ce qui expliquerait les intensités de baisse différentes.

GLM Tweedie (inc.pr.) :

- Les provisions Best Estimate sont proches du résultat de la méthode du Bootstrap à un an utilisée (-953 k€).
- L'écart-type augmente de manière significative par rapport à l'écart-type Central (+7734 k€).

La différence significative entre l'écart-type Central et l'écart-type du GLM Tweedie peut être expliquée pour les mêmes raisons précédentes.

Pour les deux LoB étudiées, le sens de variations de provisions Best Estimate et de l'écart-type par rapport aux résultats Central est identique à l'exception des provisions Best Estimate obtenues par application du GLM Tweedie :

med.exp.	GLM Poisson	GLM Tweedie	inc.pr.	GLM Poisson	GLM Tweedie
Provisions BE (k€)	0,06%	-0,04%	Provisions BE (k€)	0,01%	0,22%
MSEP (k€)	-58,23%	144,73%	MSEP (k€)	-13,08%	44,58%

Figure 43 – Variation de provisions BE et écart-type à un an par rapport aux résultats du Bootstrap à un an

L'écart entre les provisions Best Estimate des GLM et les provisions Best Estimate du modèle représente se trouve dans l'intervalle $\pm 3\%$ pour les deux LoB étudiées. Les provisions Best Estimate estimées par GLM restent proche des provisions Best Estimate estimées du modèle utilisé.

Il est également intéressant de se pencher vers la distribution des boni-mali estimés en vision un an obtenus lors de l'application de la méthode du Bootstrap. Pour compléter la représentation graphique des distributions, nous nous intéressons également à certaines grandeurs représentatives des simulations obtenues : minimum, maximum, étendue et moyenne.

Les distributions de boni-mali de la LoB **medical expenses** sont présentées ci-dessous :

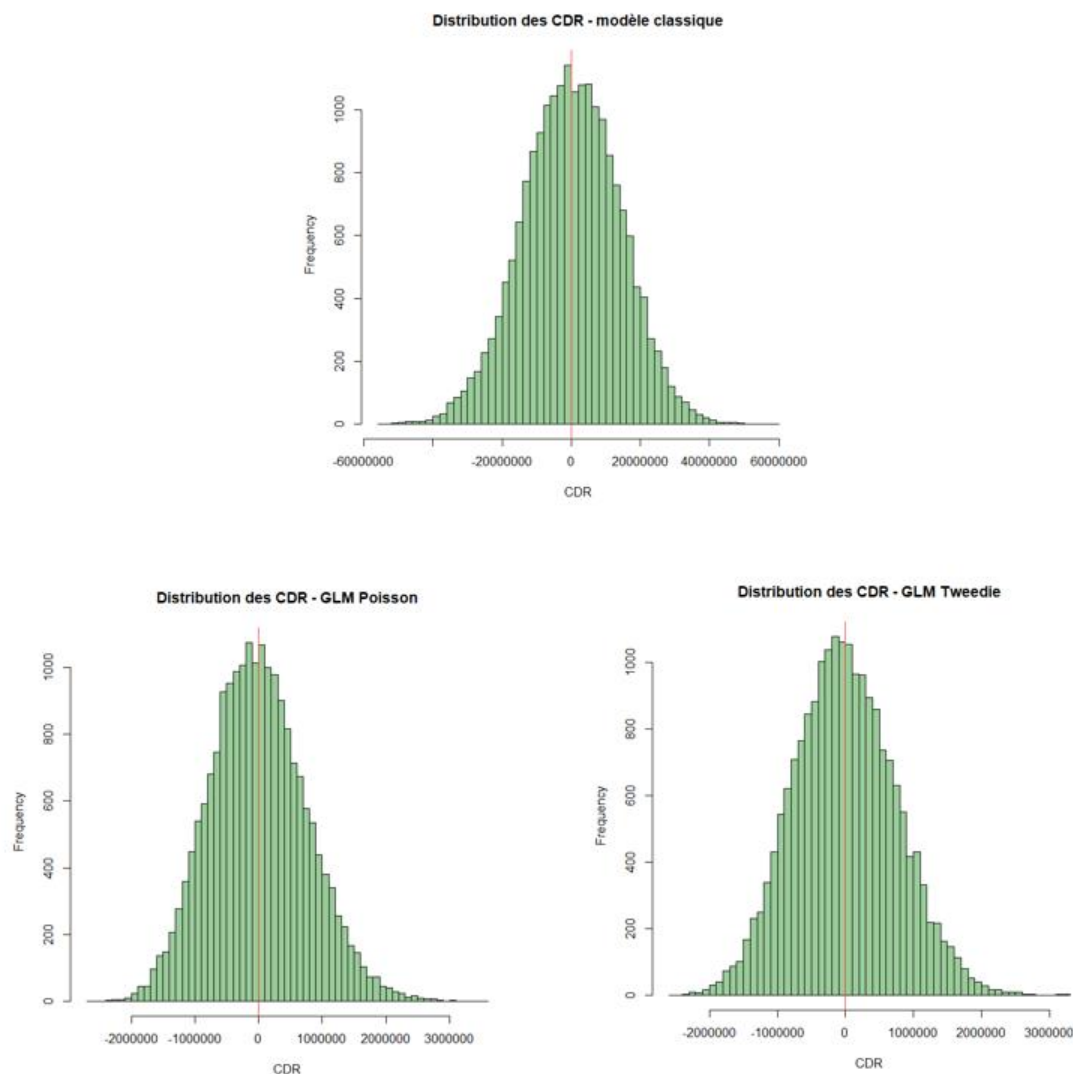


Figure 44 – Distributions des CDR (medical expenses)

La droite verticale rouge représente la moyenne des boni-mali. Pour chacune des méthodes, les boni-malis sont centrés autour de la moyenne ce qui est vrai en théorie : les provisions estimées en année N pour une année $N + 1$ devraient tendre vers les paiements effectués durant l'année $N + 1$.

med.exp.	Central	GLM Poisson	GLM Tweedie
min	-54 132 819	-2 659 087	-2 521 735
max	58 546 206	3 537 894	3 216 273
étendue	112 679 025	6 196 981	5 738 008
moyenne	4 501	-4 079	-8 322
déviance	-	30 412 658	51 434

Figure 45 – Minimum, maximum, étendue et moyenne des CDR estimés (medical expenses)

L'étendue des 20000 simulations réalisées après l'application du GLM Tweedie est plus faible que l'étendue des valeurs obtenues pour le GLM Poisson. Théoriquement, la loi Poisson est moins dispersée que la loi Tweedie telle que la puissance de variance $1 < p < 2$. En effet, la variance d'une loi Poisson est égale à son espérance : $Var(X) = E(X) = \mu$ tandis que la variance d'une loi Tweedie est égale à son espérance élevée à la puissance p : $Var(Y) = E(Y) = \mu^p$. Par conséquent, $Var(X) < Var(Y)$ ce qui implique que les simulations de loi Tweedie devraient être plus dispersées que les simulations de loi Poisson. Cependant, nos résultats montrent que le GLM Poisson est plus dispersé que le GLM Tweedie.

Les distributions de boni-mali de la **LoB income protection** sont présentées à leur tour :

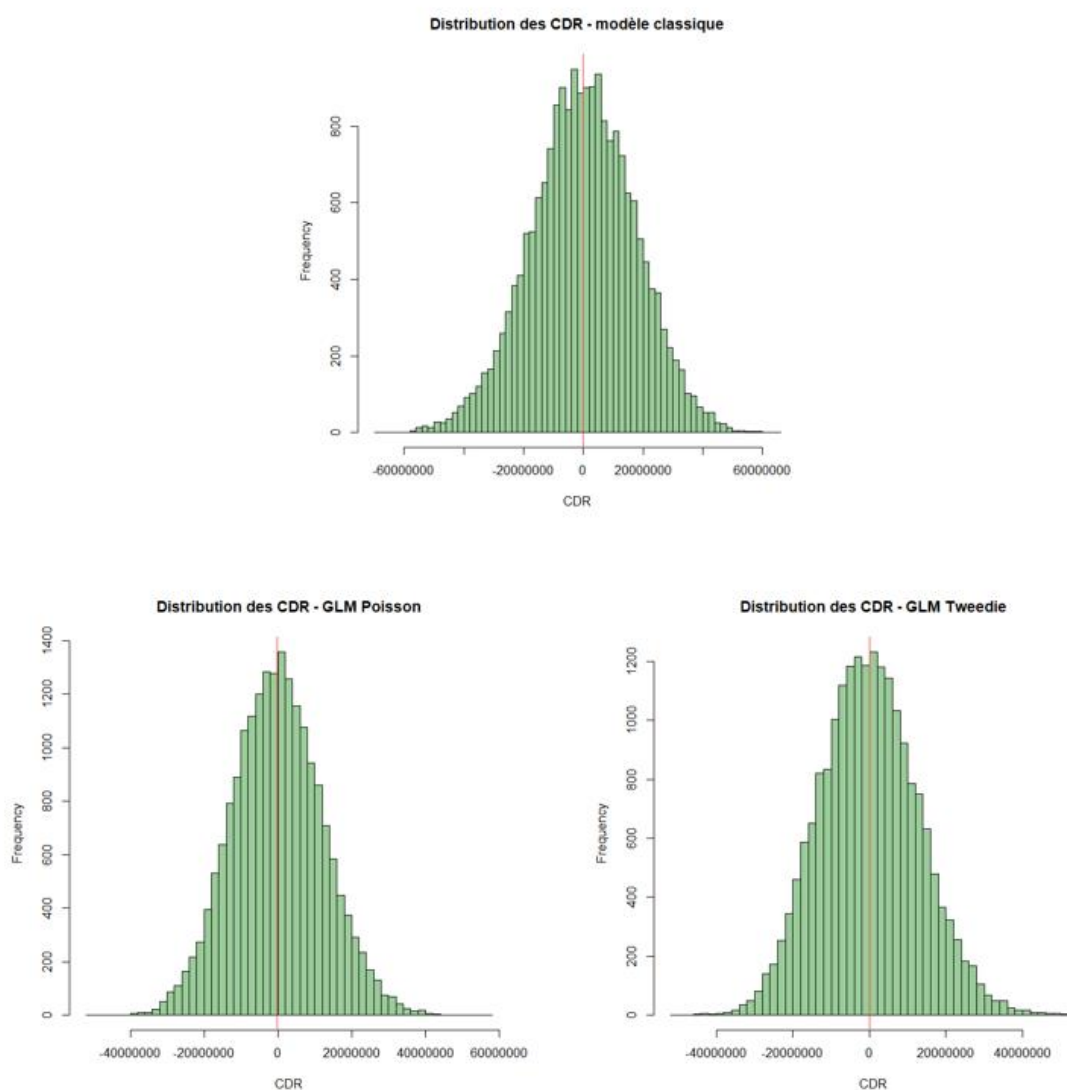


Figure 46 – Distributions des CDR (income protection)

La droite verticale rouge représente la moyenne des boni-mali. Pour chacune des méthodes, les boni-malis sont centrés autour de la moyenne ce qui est vrai en théorie : les provisions estimées en année N pour une année $N + 1$ devraient tendre vers les paiements effectués durant l'année $N + 1$.

inc.pr.	Central	GLM Poisson	GLM Tweedie
min	-68 179 611	-51 933 591	-51 276 392
max	64 438 119	56 910 137	53 326 887
étendue	132 617 730	108 843 728	104 603 279
moyenne	-42 832	-68 885	-38 716

Figure 47 – Minimum, maximum, étendue et moyenne des CDR estimés (income protection)

L'étendue des 20000 simulations réalisées après l'application du GLM Tweedie est plus faible que l'étendue des valeurs obtenues pour le GLM Poisson. Théoriquement, la loi Poisson est moins dispersée que la loi Tweedie telle que la puissance de variance $1 < p < 2$. En effet, la variance d'une loi Poisson est égale à son espérance : $Var(X) = E(X) = \mu$ tandis que la variance d'une loi Tweedie est égale à son espérance élevée à la puissance p : $Var(Y) = E(Y) = \mu^p$. Par conséquent, $Var(X) < Var(Y)$ ce qui implique que les simulations de loi Tweedie devraient être plus dispersées que les simulations de loi Poisson. Cependant, nos résultats montrent que le GLM Poisson est plus dispersé que le GLM Tweedie.

Certains indicateurs clés sont calculés afin de pousser l'analyse des résultats. En notant $(y)_{i \in \{1, \dots, N\}}$ les observations réelles et $(\hat{y})_{i \in \{1, \dots, N\}}$ les estimations réalisées par le modèle, les indicateurs utilisés sont les suivants :

- SSE : Sum of Squared Error (Somme des Carrés des Résidus SCR).

$$SSE = \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Cet indicateur majore l'importance des grosses erreurs en élevant le terme d'erreur au carré.

- MSE : Mean Squared Error (Moyenne des Carrés des Erreurs MCE)

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Cette moyenne représente la variance résiduelle du modèle. Le modèle est d'autant meilleur que la MSE est faible.

- RMSE : Rooted Mean Squared Error (Racine carrée de la Moyenne des Carrés des Erreurs RMCE)

$$RMSE = \sqrt{MSE}$$

- MAE : Mean Absolute Error (Erreur Absolue Moyenne EAM)

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

- MAPE : Mean Absolute Percentage Error (Erreur Absolue Moyenne en Pourcentage)

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{y_i}$$

Cet indicateur s'applique uniquement aux valeurs strictement positives.

L'ensemble des indicateurs présentés ci-dessus sont calculés sur les données cumulées des triangles des paiements afin d'avoir uniquement des valeurs strictement positives.

L'analyse des indicateurs de la **LoB medical expenses** est effectuée ci-dessous.

Indicateurs	Central	GLM Poisson	GLM Tweedie
SSE (10 ¹²)	2 185	2 184	134 586
MSE (10 ⁹)	6 296	6 293	385 806
RMSE (10 ³)	2 509	2 509	19 642
MAE (10 ³)	478	464	17 528
MAPE	0,075%	0,073%	2,783%

Figure 48 – Indicateurs de qualité des modèles (medical expenses)

L'ensemble des indicateurs étudiés mènent à conclure que le modèle GLM Tweedie est le modèle de moins bonne qualité. Tous ses indicateurs sont très supérieurs aux indicateurs mesurés sur les deux autres modèles.

Le modèle GLM Poisson est assez proche du modèle du Bootstrap à un an utilisé. Cela contredit les premières conclusions sur la qualité de l'estimation des paramètres du modèles. En effet, lors de l'analyse de la distribution des CDR, les valeurs minimales, maximales et moyennes des boni-mali étaient significativement différentes des résultats obtenus par application du Bootstrap à un an.

L'analyse des indicateurs de la **LoB income protection** est effectuée à son tour.

Indicateurs	Central	GLM Poisson	GLM Tweedie
SSE (10 ¹²)	3 392	3 400	25 998
MSE (10 ⁹)	9 776	9 799	74 300
RMSE (10 ³)	3 127	3 130	8 620
MAE (10 ³)	1 622	1 609	7 010
MAPE	1,244%	1,236%	4,301%

Figure 49 – Indicateurs de qualité des modèles (income protection)

De manière similaire à la LoB medical expenses, l'ensemble des indicateurs étudiés mènent à conclure que le modèle GLM Tweedie est le modèle de moins bonne qualité.

Le modèle GLM Poisson est assez proche du modèle du Bootstrap à un an. Le nombre de valeurs négatives dans les données de cette LoB étant moins important que la LoB medical expenses, ce résultat est justifié par l'équivalence entre la méthode Chain-Ladder et le modèle GLM Poisson.

Un dernier indicateur est intéressant à mesurer : l'indice de Gini. Cet indice est une mesure statistique permettant d'analyser la concentration d'une variable au sein d'une population donnée. Il est compris entre 0 et 1 et mesure ainsi l'inégalité de la répartition de la variable étudiée. Soit X la variable étudiée :

- $Gini_X = 0$: répartition parfaitement égalitaire
- $Gini_X = 1$: répartition parfaitement inégalitaire

L'indice de Gini est associé à la courbe de Lorenz permettant de visualiser la concentration de la variable étudiée. Cette représentation graphique de la concentration de la variable permet de calculer l'indice de Gini par un calcul d'aire. La courbe de Lorenz est comparée à la droite de distribution égalitaire de la variable.

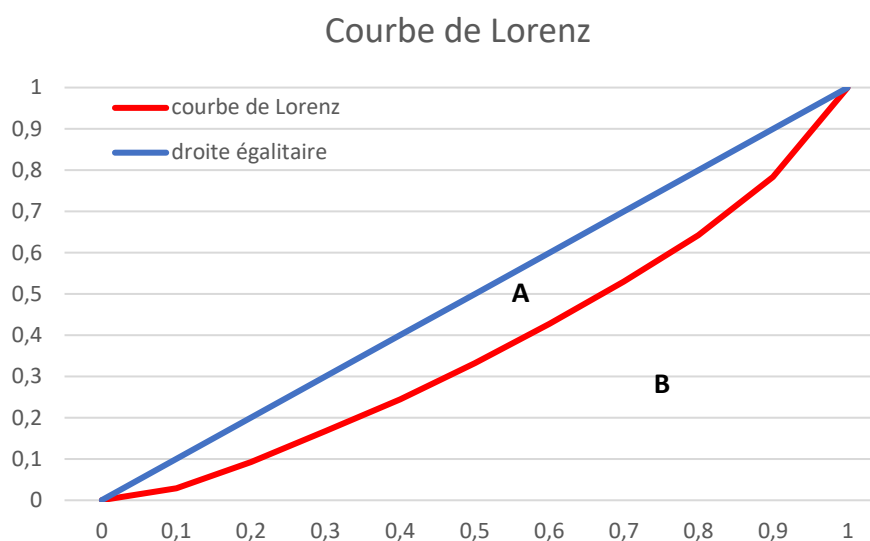


Figure 50 – Courbe de Lorenz

Sur le graphique ci-dessus, l'aire située entre la droite égalitaire et la courbe de Lorenz est la surface de concentration (notée A). L'aire sous la courbe de Lorenz est (notée B). Le coefficient de Gini se calcule alors de la manière suivante :

$$Gini_X = \frac{A}{A + B}$$

Dans notre étude, l'indice de Gini permet d'observer la dispersion des réserves par rapport aux années de survenance estimées dans les différents modèles. Nous pourrions ainsi identifier quelles sont les années

de survie contribuant le plus aux provisions Best Estimate. Les indices de Gini ont été mesurés à l'aide de la fonction `ineq()` incluse dans le package `ineq` du logiciel R.

Les résultats obtenus pour la LoB medical expenses sont présentés ci-dessous.

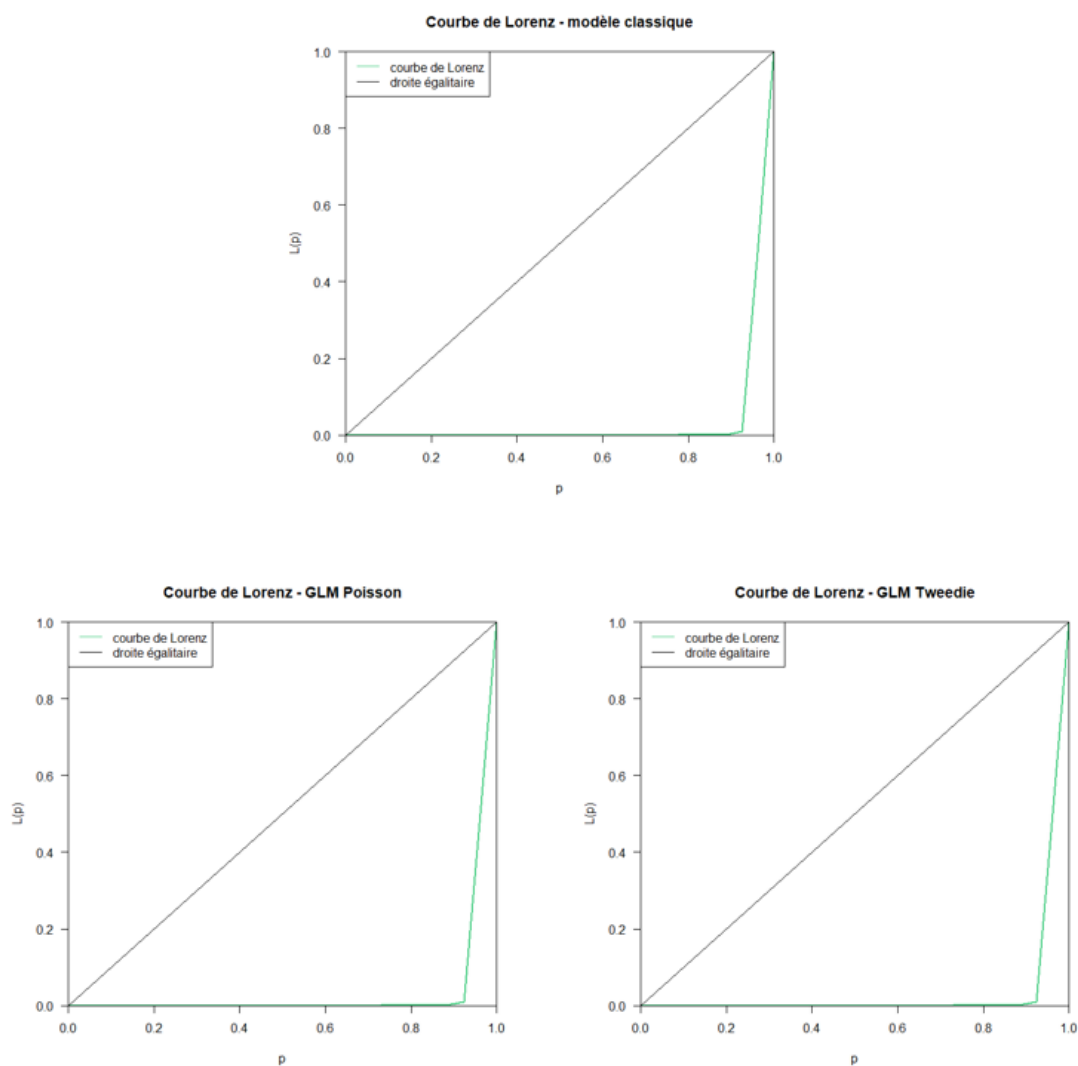


Figure 51 – Courbes de Lorenz des réserves en fonction des années de survie (medical expenses)

med.exp.	Central	GLM Poisson	GLM Tweedie
indice de Gini	0,9250480	0,9221151	0,9221151

Figure 52 – Indices de Gini des réserves (medical expenses)

Les modèles étudiés aboutissent au même résultat. La répartition des réserves est inégalitaire et sont concentrées sur les années de survie les plus récentes. Cela remet en question la profondeur d'historique utilisée au sein du modèle.

Les résultats obtenus pour la LoB income protection sont présentés à leur tour.

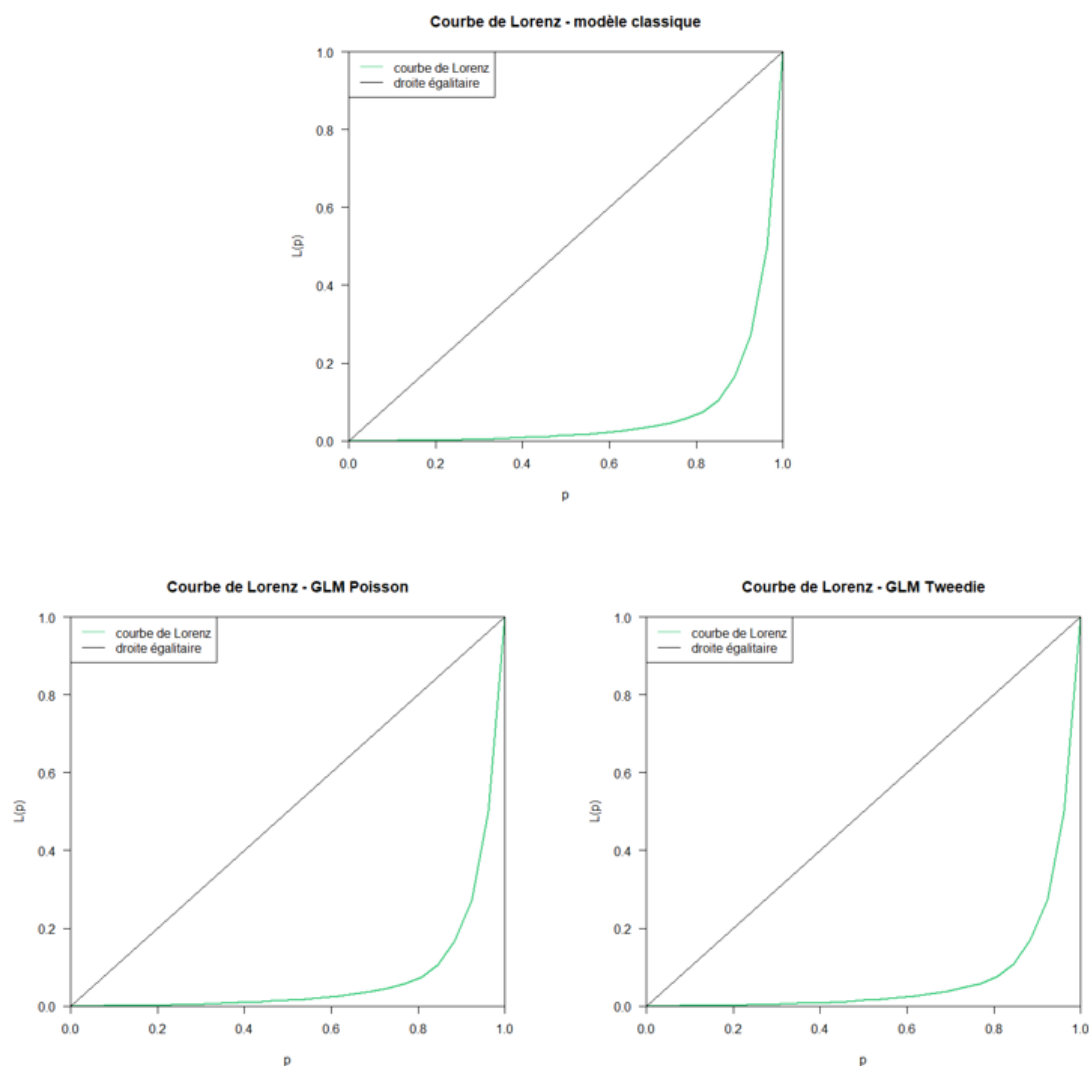


Figure 53 – Courbes de Lorenz des réserves en fonction des années de survivance (income protection)

inc.pr.	Central	GLM Poisson	GLM Tweedie
indice de Gini	0,8575660	0,8520204	0,8508477

Figure 54 – Indices de Gini des réserves (income protection)

Les résultats obtenus sont similaires à la LoB medical expenses. Les indices de Gini sont moins élevés mais la répartition des réserves reste inégalitaire. La concentration des réserves par années concerne également certaines années de survivance. La profondeur d'historique lors du provisionnement est à étudier LoB.

2. Limites

Comme tous modèles, les GLM possèdent certaines limites :

- Les paiements incrémentaux doivent être positifs pour l'application des modèles. Cependant, il est possible de rencontrer des montants négatifs dans les triangles de paiements incrémentaux suite à des recours. Les incréments négatifs peuvent être remplacés par des valeurs nulles ou directement supprimés, ce qui conduit à un modèle de moins bonne qualité. Cependant, l'application d'un GLM sur le triangle des facteurs de développement individuels permettrait de résoudre le problème des incréments négatifs. En effet, les coefficients de développement individuels sont compris entre 0 et 1 et à partir de ce triangle, il est possible de reconstruire le triangle des paiements.
- Le nombre de paramètres à estimer est élevé. Par conséquent, une bonne estimation nécessite un nombre important de données.

Après l'analyse et l'étude des résultats obtenus, le passage de la méthode du Bootstrap à un an vers l'application d'un GLM Tweedie pour le calcul de volatilité à un an n'aboutit pas à une conclusion favorable. L'estimation de volatilité à un an par GLM Tweedie est plus élevée que celle obtenue par Bootstrap à un an, ce qui sous-entend que le GLM Tweedie tient compte d'une corrélation implicite entre années de survenance plus élevée et permet d'être plus prudent au niveau du calcul du SCR de risque de réserves. Cependant, la qualité d'estimation remet en cause son utilisation mais se retrouve également limitée par la nature des données incrémentales.

CONCLUSION

L'étude menée tout au long de ce mémoire consiste à étudier les possibilités d'intégrer aux mieux la prise en compte de la dépendance entre années de survénance au calcul de volatilité à un an du risque de réserves dans le cas où les tests statistiques mis en place par Groupama sont rejetés.

Au sein du modèle interne de Groupama, le calcul du SCR de risque de réserves est réalisé à l'aide de la méthode Merz-Wuthrich Chain-Ladder. La méthode Chain-Ladder repose sur trois hypothèses dont l'hypothèse d'indépendance entre années de survénance. Pour certaines lignes d'activité, seule cette hypothèse est rejetée comme nous avons pu le voir pour les deux lignes d'activités étudiées (medical expenses et income protection)

Dans un premier temps, nous nous sommes intéressés à un facteur calendaire pouvant être à l'origine du rejet du test d'indépendance entre années de survénance : l'inflation. Les résultats obtenus nous amènent à conclure que l'inflation n'est pas à l'origine de la présence d'effets calendaires au sein des triangles des données.

Ensuite, nous avons testé l'ajout de dépendance selon deux méthodes. La première méthode consiste à agréger les volatilités à un an individuelles (i.e. par années de survénance) à l'aide d'une matrice d'équi-corrélation afin de calculer la volatilité à un an totale. La seconde méthode consiste à créer de la dépendance par ajout d'une copule gaussienne lors de la simulation de la diagonale $N+1$. Les résultats obtenus ne permettent pas de conserver l'une ou l'autre méthode. Cependant, l'étude du comportement des années de survénance entre elles nous amène à penser qu'une méthode de corrélation par bloc semble opportune. Cette idée se rapproche de la méthode d'agrégation par matrice d'équi-corrélation. La matrice de corrélation serait alors composée de plusieurs sous matrice d'équi-corrélation, dont le coefficient de corrélation de chaque sous-matrice est adapté au degré de liaison observé.

Pour finir, nous avons appliqué la méthode de provisionnement par utilisation de modèles linéaires généralisés en substitution de la méthode Chain-Ladder. Cette étude est réalisée afin d'analyser le comportement de la dépendance lors d'un changement de modèle restant similaire à la méthode actuelle. En effet, le GLM Poisson et la méthode Chain-Ladder sont supposées converger vers les mêmes résultats. Cependant, les paiements incrémentaux présentent des valeurs négatives. L'alternative était de remplacer les valeurs négatives par des valeurs nulles mais cela entraîne une baisse de qualité d'estimation qui entraîne la divergence des modèles sur certains résultats. Nous avons également appliqué un GLM Tweedie qui regroupe l'ensemble des familles de lois exponentielles. Les résultats obtenus ne nous ont pas permis de conserver l'utilisation du GLM Tweedie lors de l'analyse de certains indicateurs.

Il est à noter que compte tenu du contexte d'inflation, les méthodes ayant recours aux facteurs de développement de Chain-Ladder sont remises en cause. En effet, ces coefficients projettent une inflation

passée constante lors de l'estimation des ultimes, ce qui ne semble pas représentatif du contexte actuel. Par conséquent, la méthode du Bootstrap à un an utilisée pour le calcul du SCR de risque de réserves au sein du modèle interne est à faire évoluer.

ANNEXES

Annexe 1 : Triangles des règlements cumulés (k€) des LoB étudiées

Le triangle des paiements cumulés (k€) de la LoB medical expenses est donné ci-dessous :

AY\DY	1	2	3	4	5	6	7	8	9	10	11	12	13
1996	379 070	469 338	469 949	469 967	469 972	469 974	469 975	469 975	469 976	469 976	469 976	469 976	469 976
1997	384 249	478 103	478 376	478 519	478 519	478 519	478 520	478 519	478 519	478 520	478 520	478 520	478 520
1998	400 438	497 269	497 710	497 746	497 750	497 754	497 754	497 754	497 754	497 755	497 756	497 754	497 756
1999	410 367	515 747	516 316	516 558	516 515	516 519	516 519	516 520	516 521	516 521	516 521	516 521	516 521
2000	433 773	543 623	544 527	544 764	544 785	545 086	545 081	545 080	545 129	545 129	545 129	545 129	545 129
2001	450 972	563 829	565 282	565 372	565 478	565 488	565 486	565 491	565 491	565 491	565 491	565 491	565 491
2002	480 574	589 015	590 096	590 338	590 363	590 390	590 391	590 394	590 394	590 394	590 395	590 395	590 393
2003	510 896	622 026	626 445	626 498	626 334	626 334	626 320	626 320	626 324	626 327	626 338	626 335	626 335
2004	539 363	656 992	658 493	658 514	658 596	658 610	658 611	658 612	658 612	658 613	658 615	658 615	658 615
2005	557 243	681 026	682 635	682 814	682 877	682 869	682 871	682 871	682 871	682 851	682 851	682 851	682 851
2006	560 663	704 623	706 138	706 225	706 230	706 240	706 242	706 247	706 354	706 354	706 354	706 354	706 353
2007	602 835	737 310	738 834	738 896	738 924	738 913	738 916	738 911	738 911	738 911	738 911	738 916	738 919
2008	666 201	824 984	826 708	826 861	826 859	826 856	826 881	826 883	826 884	826 884	826 885	826 885	826 885
2009	705 430	873 730	876 296	876 437	876 480	876 424	876 434	876 440	876 444	876 450	876 452	876 453	876 453
2010	752 938	934 496	936 830	936 985	936 939	936 949	936 960	936 966	936 968	936 968	936 967	936 970	
2011	814 547	1 007 727	1 010 154	1 010 429	1 010 454	1 010 467	1 010 475	1 010 484	1 010 493	1 010 488	1 010 490		
2012	825 852	1 010 056	1 012 419	1 012 658	1 012 703	1 012 727	1 012 741	1 012 755	1 012 758	1 012 759			
2013	807 295	996 185	999 362	999 530	999 622	999 659	999 692	999 695	999 739				
2014	792 166	984 701	987 569	987 892	988 118	988 212	988 217	988 253					
2015	782 619	971 125	974 417	974 860	975 154	975 218	975 251						
2016	769 155	951 779	954 646	955 026	955 198	955 227							
2017	773 509	951 079	954 238	954 596	954 763								
2018	801 371	984 277	986 875	988 016									
2019	814 910	1 009 044	1 012 952										
2020	771 116	1 003 842											
2021	910 762												

AY\DY	14	15	16	17	18	19	20	21	22	23	24	25	26
1996	469 976	469 976	469 976	469 976	469 976	469 976	469 976	469 976	469 976	469 976	469 976	469 976	469 976
1997	478 521	478 521	478 521	478 521	478 521	478 521	478 521	478 521	478 521	478 521	478 521	478 521	478 521
1998	497 756	497 757	497 757	497 757	497 758	497 758	497 758	497 758	497 758	497 758	497 758		
1999	516 521	516 521	516 521	516 521	516 521	516 521	516 521	516 521	516 521	516 521			
2000	545 129	545 129	545 129	545 129	545 129	545 129	545 129	545 129	545 129				
2001	565 490	565 490	565 490	565 490	565 490	565 490	565 490	565 490					
2002	590 393	590 393	590 393	590 393	590 393	590 393	590 393						
2003	626 335	626 335	626 342	626 342	626 342								
2004	658 615	658 615	658 615	658 615	658 615								
2005	682 851	682 851	682 851	682 851									
2006	706 355	706 355	706 355										
2007	738 919	738 919											
2008	826 885												
2009													
2010													
2011													
2012													
2013													
2014													
2015													
2016													
2017													
2018													
2019													
2020													
2021													

Le triangle des règlements cumulés (k€) de la LoB income protection est donné ci-dessous :

AY\DY	1	2	3	4	5	6	7	8	9	10	11	12	13
1996	60 018	130 048	148 054	157 561	161 891	164 018	164 775	165 613	165 424	165 773	166 141	166 481	166 852
1997	59 144	130 400	148 583	158 348	163 168	165 660	167 009	167 499	167 818	168 531	169 312	170 010	170 452
1998	60 290	136 260	155 525	168 338	174 677	177 294	178 936	179 838	180 858	181 320	181 374	181 700	181 829
1999	60 420	136 115	156 799	167 725	173 568	176 219	177 790	178 834	179 794	180 288	180 973	181 090	181 395
2000	62 771	141 274	164 458	177 749	184 120	187 127	188 987	190 213	191 125	191 376	191 800	192 181	192 569
2001	62 667	144 502	167 699	181 773	192 407	195 298	198 104	199 041	199 722	200 212	200 745	201 288	201 577
2002	60 299	140 005	162 942	175 653	181 965	184 319	185 560	186 397	187 863	188 157	188 315	188 571	188 657
2003	56 991	129 129	150 466	163 597	169 562	171 388	172 630	173 371	174 492	174 942	175 368	175 602	175 945
2004	57 001	131 553	152 128	164 579	171 573	174 274	175 979	176 553	176 997	177 247	177 500	177 638	177 658
2005	60 044	138 543	162 724	175 436	180 829	183 406	185 157	186 093	186 773	187 121	187 703	187 914	187 997
2006	58 360	138 414	161 844	175 554	181 694	185 767	188 658	189 956	190 863	191 254	191 618	192 105	192 271
2007	66 556	136 391	161 348	176 398	182 255	185 305	186 989	188 106	188 867	189 503	189 860	190 160	190 414
2008	63 475	138 398	165 803	180 568	187 527	190 767	192 552	193 789	194 380	194 791	194 978	195 223	195 852
2009	68 263	147 969	176 792	192 122	199 438	202 477	204 469	205 489	207 810	208 303	208 584	208 894	209 300
2010	71 961	157 197	187 879	205 846	213 028	216 253	218 663	219 684	220 246	221 378	221 763	222 107	
2011	72 122	156 358	188 471	207 017	214 876	219 718	222 961	224 598	225 763	225 900	226 331		
2012	72 621	158 034	191 511	213 597	223 193	227 680	230 853	231 937	232 853	233 537			
2013	72 878	162 270	197 323	221 039	232 825	236 650	238 421	239 411	240 280				
2014	70 914	162 004	197 754	219 735	229 297	233 938	237 828	239 913					
2015	70 464	166 661	205 189	227 740	238 660	243 068	244 890						
2016	73 599	167 051	207 426	230 826	239 856	248 494							
2017	78 096	185 872	236 029	258 370	271 906								
2018	78 349	188 381	235 837	269 828									
2019	85 645	204 983	263 649										
2020	83 819	209 694											
2021	88 546												

AY\DY	14	15	16	17	18	19	20	21	22	23	24	25	26
1996	167 098	167 262	167 390	167 516	167 652	167 742	167 868	168 035	168 124	168 176	168 268	168 348	168 389
1997	171 092	171 557	171 829	172 156	172 441	172 788	173 059	173 381	173 614	173 930	174 149	174 425	
1998	182 217	182 305	182 658	183 042	183 151	183 338	183 533	183 867	183 979	184 042	184 103		
1999	181 867	182 138	182 275	182 382	182 489	182 678	182 769	182 861	182 943	183 020			
2000	192 678	192 877	193 208	193 297	193 603	193 769	193 856	193 920	193 996				
2001	201 599	201 770	201 962	202 049	202 241	202 392	202 484	202 601					
2002	188 828	188 892	188 967	189 004	189 035	189 074	189 099						
2003	176 110	176 285	176 476	176 642	176 753	176 779							
2004	177 933	178 022	178 182	178 270	178 332								
2005	188 041	188 150	188 224	188 302									
2006	192 424	192 658	193 175										
2007	190 685	190 784											
2008	196 145												
2009													
2010													
2011													
2012													
2013													
2014													
2015													
2016													
2017													
2018													
2019													
2020													
2021													

Annexe 2 : Triangles des règlements incrémentaux (k€) des LoB étudiées

Le triangle des règlements incrémentaux (k€) de la LoB medical expenses est donné ci-dessous :

AY\DY	1	2	3	4	5	6	7	8	9	10	11	12	13
1996	379 070	90 268	611	18	5	2	1	0	1	0	0	0	0
1997	384 249	93 854	273	144	0	0	1	-1	0	1	0	0	0
1998	400 438	96 831	441	36	3	4	1	-1	0	1	0	-2	2
1999	410 367	105 380	568	242	-42	4	0	1	1	0	0	0	0
2000	433 773	109 850	904	237	21	301	-5	-1	49	0	0	0	0
2001	450 972	112 857	1 453	90	105	11	-3	5	0	0	0	0	0
2002	480 574	108 441	1 081	242	25	26	1	3	0	0	0	0	-1
2003	510 896	111 129	4 419	53	-165	1	-15	0	4	3	11	-3	0
2004	539 363	117 629	1 500	22	81	14	1	1	0	3	0	0	0
2005	557 243	123 783	1 609	180	62	-7	1	0	0	-20	0	0	0
2006	560 663	143 960	1 515	87	5	10	3	4	107	0	0	-1	0
2007	602 835	134 475	1 524	62	28	-10	2	-5	0	0	0	5	3
2008	666 201	158 783	1 725	152	-2	-3	25	3	1	1	0	0	0
2009	705 430	168 300	2 566	142	42	-56	10	6	4	5	2	1	1
2010	752 938	181 558	2 334	154	-45	10	11	6	2	0	0	3	
2011	814 547	193 180	2 427	275	25	13	8	9	9	-4	2		
2012	825 852	184 203	2 364	238	45	24	15	13	3	1			
2013	807 295	188 890	3 177	168	92	37	33	3	45				
2014	792 166	192 535	2 868	323	226	94	5	36					
2015	782 619	188 507	3 291	443	294	64	32						
2016	769 155	182 625	2 867	380	172	29							
2017	773 509	177 570	3 159	358	167								
2018	801 371	182 906	2 598	1 141									
2019	814 910	194 134	3 908										
2020	771 116	232 726											
2021	910 762												

AY\DY	14	15	16	17	18	19	20	21	22	23	24	25	26
1996	0	0	0	0	0	0	0	0	0	0	0	0	0
1997	1	0	0	0	0	0	0	0	0	0	0	0	0
1998	1	1	0	0	1	1	0	0	0	0	0	0	0
1999	0	0	1	0	0	0	0	0	0	0	0	0	0
2000	0	0	0	0	0	0	0	0	0	0	0	0	0
2001	-1	0	0	0	0	0	0	0	0	0	0	0	0
2002	0	0	0	0	0	0	0	0	0	0	0	0	0
2003	0	0	6	0	0	0	0	0	0	0	0	0	0
2004	0	0	0	0	0	0	0	0	0	0	0	0	0
2005	0	0	0	0	0	0	0	0	0	0	0	0	0
2006	2	0	0	0	0	0	0	0	0	0	0	0	0
2007	0	0	0	0	0	0	0	0	0	0	0	0	0
2008	0	0	0	0	0	0	0	0	0	0	0	0	0
2009	0	0	0	0	0	0	0	0	0	0	0	0	0
2010	0	0	0	0	0	0	0	0	0	0	0	0	0
2011	0	0	0	0	0	0	0	0	0	0	0	0	0
2012	0	0	0	0	0	0	0	0	0	0	0	0	0
2013	0	0	0	0	0	0	0	0	0	0	0	0	0
2014	0	0	0	0	0	0	0	0	0	0	0	0	0
2015	0	0	0	0	0	0	0	0	0	0	0	0	0
2016	0	0	0	0	0	0	0	0	0	0	0	0	0
2017	0	0	0	0	0	0	0	0	0	0	0	0	0
2018	0	0	0	0	0	0	0	0	0	0	0	0	0
2019	0	0	0	0	0	0	0	0	0	0	0	0	0
2020	0	0	0	0	0	0	0	0	0	0	0	0	0
2021	0	0	0	0	0	0	0	0	0	0	0	0	0

Le triangle des règlements incrémentaux (k€) de la LoB income protection est donné ci-dessous :

AY\DY	1	2	3	4	5	6	7	8	9	10	11	12	13
1996	60 018	70 031	18 005	9 507	4 331	2 127	757	838	-188	349	368	340	371
1997	59 144	71 256	18 183	9 764	4 821	2 492	1 349	489	320	713	781	698	442
1998	60 290	75 970	19 265	12 813	6 339	2 618	1 641	902	1 020	462	54	326	129
1999	60 420	75 695	20 684	10 925	5 843	2 651	1 571	1 044	959	494	685	117	306
2000	62 771	78 504	23 184	13 291	6 371	3 007	1 860	1 226	912	252	424	381	388
2001	62 667	81 835	23 197	14 074	10 634	2 891	2 806	937	680	490	533	543	288
2002	60 299	79 706	22 937	12 711	6 312	2 354	1 241	837	1 466	294	158	255	87
2003	56 991	72 138	21 337	13 131	5 965	1 826	1 242	741	1 121	450	426	234	343
2004	57 001	74 552	20 575	12 631	6 813	2 702	1 705	574	444	250	253	138	20
2005	60 044	78 499	24 180	12 712	5 393	2 577	1 751	936	681	347	582	211	83
2006	58 360	80 053	23 430	13 711	6 140	4 073	2 891	1 298	907	390	364	487	166
2007	66 556	69 836	24 956	15 051	5 857	3 050	1 684	1 117	762	635	357	300	254
2008	63 475	74 923	27 405	14 765	6 959	3 239	1 786	1 237	590	412	186	245	629
2009	68 263	79 706	28 823	15 331	7 315	3 039	1 992	1 020	2 321	493	281	311	405
2010	71 961	85 235	30 682	17 967	7 182	3 225	2 410	1 021	563	1 132	385	344	
2011	72 122	84 236	32 113	18 546	7 859	4 842	3 243	1 637	1 165	138	431		
2012	72 621	85 413	33 477	22 086	9 596	4 486	3 173	1 084	917	683			
2013	72 878	89 391	35 053	23 716	11 786	3 824	1 771	990	868				
2014	70 914	91 090	35 750	21 982	9 562	4 640	3 890	2 086					
2015	70 464	96 197	38 529	22 550	10 920	4 408	1 821						
2016	73 599	93 452	40 375	23 400	9 030	8 639							
2017	78 096	107 777	50 157	22 340	13 537								
2018	78 349	110 032	47 456	33 991									
2019	85 645	119 339	58 666										
2020	83 819	125 876											
2021	88 546												

AY\DY	14	15	16	17	18	19	20	21	22	23	24	25	26
1996	247	164	127	127	136	90	126	167	89	52	92	79	41
1997	640	465	272	327	285	348	271	322	233	316	219	276	
1998	388	89	353	384	109	187	195	334	112	62	62		
1999	472	270	138	106	108	188	91	92	82	77			
2000	109	198	331	90	306	166	86	64	76				
2001	23	171	192	87	193	150	92	117					
2002	170	65	75	36	32	39	24						
2003	165	176	191	166	110	26							
2004	275	89	161	88	62								
2005	44	109	74	78									
2006	153	234	517										
2007	272	99											
2008	293												
2009													
2010													
2011													
2012													
2013													
2014													
2015													
2016													
2017													
2018													
2019													
2020													
2021													

Dans les deux triangles présentés, les cellules rouges correspondent aux règlements négatifs.

Annexe 3 : Triangles des charges cumulées (k€) de la LoB income protection

AY\DY	1	2	3	4	5	6	7	8	9	10	11	12	13
1996	134 189	189 978	181 268	177 201	175 998	173 982	173 118	171 375	170 013	170 012	169 965	170 041	169 993
1997	135 172	191 906	186 229	183 096	178 829	177 877	177 596	177 419	176 953	177 368	177 102	177 468	178 212
1998	138 510	201 278	197 866	192 961	191 183	191 186	190 271	189 888	189 263	189 196	187 628	187 396	186 673
1999	130 174	194 559	193 900	188 050	187 254	186 752	185 269	185 370	185 755	185 430	185 158	184 495	184 305
2000	118 040	208 216	205 568	204 255	202 260	200 577	198 796	199 010	198 472	197 448	196 885	196 600	195 829
2001	150 388	214 925	212 603	210 300	212 371	210 435	210 869	210 480	209 679	208 281	206 995	205 927	205 657
2002	146 901	199 437	196 142	194 623	192 814	192 862	191 949	190 711	190 430	189 932	189 570	189 696	189 575
2003	139 395	183 740	183 392	179 492	179 517	178 147	177 338	177 167	176 458	176 404	176 785	176 747	176 903
2004	139 486	188 893	185 103	184 309	182 555	181 409	179 968	178 526	178 559	178 542	178 375	178 354	178 187
2005	144 196	197 095	195 491	193 642	191 321	190 038	189 159	189 133	189 163	189 241	189 118	188 990	188 486
2006	133 495	204 559	200 179	196 934	195 856	193 521	194 308	194 658	194 610	194 269	193 952	193 457	193 236
2007	160 523	203 459	200 736	197 292	195 532	192 735	192 121	192 142	191 866	192 040	190 678	190 741	190 936
2008	154 000	207 163	208 414	204 749	203 755	198 059	197 582	197 624	197 397	197 411	198 902	198 246	198 590
2009	155 072	212 409	216 802	213 526	210 136	210 830	210 875	210 064	210 031	209 465	209 390	209 613	210 040
2010	161 367	229 451	230 250	228 647	226 692	224 910	224 169	223 054	221 941	222 448	222 686	222 692	
2011	168 549	234 135	236 434	236 999	235 011	232 497	231 932	231 471	229 761	229 397	229 043		
2012	169 530	241 885	244 523	240 964	240 293	237 679	234 669	234 445	234 707	235 546			
2013	166 017	246 559	255 412	252 446	250 674	246 185	245 822	244 841	244 235				
2014	178 183	251 825	254 768	253 578	246 548	245 967	243 331	243 504					
2015	165 490	263 728	266 519	261 199	255 596	254 330	254 130						
2016	173 551	261 775	267 025	263 437	261 576	258 399							
2017	204 842	298 651	297 086	295 379	294 886								
2018	228 976	311 991	318 285	324 546									
2019	230 969	342 127	353 772										
2020	237 563	350 491											
2021	263 668												

AY\DY	14	15	16	17	18	19	20	21	22	23	24	25	26
1996	170 141	170 229	169 841	169 812	169 667	169 746	169 779	169 808	169 674	169 440	169 493	169 737	169 645
1997	178 477	178 222	177 953	177 957	178 160	178 522	178 603	178 585	178 291	178 438	178 604	178 554	
1998	186 451	185 843	185 963	186 104	186 002	185 841	185 529	185 754	185 675	185 686	185 652		
1999	184 288	184 271	184 280	184 033	183 944	183 859	183 893	183 805	183 825	184 030			
2000	195 466	194 987	195 165	195 236	195 186	195 159	195 036	195 063	195 064				
2001	205 351	205 237	204 994	204 409	204 504	204 514	204 570	204 676					
2002	189 502	189 349	189 368	189 392	189 341	189 393	189 392						
2003	176 851	176 935	176 780	176 933	177 019	176 947							
2004	178 405	178 306	178 363	178 382	178 400								
2005	188 450	188 393	188 448	188 471									
2006	193 425	193 579	193 837										
2007	191 131	190 932											
2008	198 449												
2009													
2010													
2011													
2012													
2013													
2014													
2015													
2016													
2017													
2018													
2019													
2020													
2021													

Annexe 4 : Triangles des règlements cumulés (k€) « as-if » 2021 des LoB étudiées

Le triangle des règlements cumulés (k€) « as-if » 2021 de la LoB medical expenses est donné ci-dessous :

AY\DY	1	2	3	4	5	6	7	8	9	10	11	12	13
1996	882 409	1 076 391	1 036 338	1 010 115	956 558	905 836	845 786	794 912	757 782	728 637	706 043	678 888	655 930
1997	881 245	1 054 319	1 028 188	973 955	922 306	861 163	809 365	771 558	741 882	718 879	691 230	667 856	644 648
1998	883 052	1 068 796	1 013 015	959 365	895 772	841 897	802 572	771 703	747 775	719 016	694 702	670 559	653 569
1999	882 015	1 049 727	995 155	929 619	873 630	832 828	800 796	775 966	746 123	720 891	695 841	678 208	660 378
2000	882 881	1 047 788	979 954	921 409	878 403	845 086	818 874	787 377	760 819	734 381	715 771	696 954	681 284
2001	869 210	1 014 690	956 114	911 598	876 700	849 532	816 854	789 238	761 812	742 507	722 987	706 732	691 519
2002	864 861	996 256	951 463	915 243	886 902	852 828	823 990	795 361	775 206	754 826	737 856	721 972	705 049
2003	864 126	1 002 945	971 222	941 187	904 749	874 155	843 759	822 377	800 762	782 762	765 926	747 971	732 587
2004	869 662	1 018 582	989 252	951 235	919 181	887 260	864 777	842 044	823 112	805 394	786 520	770 343	756 722
2005	863 934	1 023 103	986 077	952 982	919 951	896 629	873 058	853 430	835 058	815 462	798 690	784 568	767 679
2006	842 282	1 017 840	985 535	951 405	927 302	902 936	882 639	863 644	843 530	826 181	811 572	794 101	777 007
2007	870 805	1 029 040	995 335	970 194	944 723	923 470	903 593	882 410	864 260	848 979	830 703	812 827	792 232
2008	929 795	1 111 394	1 085 494	1 057 151	1 033 381	1 011 132	987 463	967 157	950 056	929 605	909 595	886 545	864 922
2009	950 334	1 147 235	1 120 354	1 095 343	1 071 816	1 046 629	1 025 113	1 006 995	985 322	964 117	939 688	916 769	876 453
2010	988 632	1 194 765	1 170 820	1 145 805	1 118 896	1 095 894	1 076 529	1 053 362	1 030 689	1 004 570	980 068	936 970	
2011	1 041 408	1 259 424	1 235 281	1 206 657	1 181 868	1 160 986	1 136 003	1 111 559	1 083 400	1 056 971	1 010 490		
2012	1 032 123	1 235 161	1 209 035	1 184 446	1 163 554	1 138 534	1 114 042	1 085 825	1 059 344	1 012 759			
2013	987 212	1 189 647	1 168 894	1 148 419	1 123 801	1 099 651	1 071 820	1 045 681	999 739				
2014	946 007	1 151 746	1 134 677	1 110 615	1 086 956	1 059 512	1 033 675	988 253					
2015	915 383	1 115 784	1 095 465	1 072 372	1 045 511	1 020 078	975 251						
2016	883 728	1 070 016	1 050 135	1 023 931	999 137	955 227							
2017	869 600	1 046 212	1 023 086	998 507	954 763								
2018	881 529	1 055 293	1 032 271	988 016									
2019	873 706	1 055 460	1 012 952										
2020	806 587	1 003 842											
2021	910 762												

AY\DY	14	15	16	17	18	19	20	21	22	23	24	25	26
1996	633 137	617 093	600 870	587 361	574 717	561 247	549 703	539 984	528 360	516 986	503 885	491 595	469 976
1997	628 313	611 795	598 040	585 166	571 451	559 698	549 801	537 966	526 385	513 046	500 533	478 521	
1998	636 387	622 080	608 689	594 423	582 198	571 904	559 593	547 547	533 672	520 655	497 758		
1999	645 530	631 634	616 831	604 144	593 462	580 687	568 187	553 788	540 281	516 521			
2000	666 619	650 995	637 605	626 331	612 848	599 656	584 460	570 205	545 129				
2001	675 310	661 421	649 725	635 739	622 054	606 290	591 503	565 490					
2002	690 548	678 338	663 736	649 448	632 990	617 551	590 393						
2003	719 634	704 142	688 992	671 532	655 153	626 342							
2004	740 433	724 494	706 134	688 911	658 615								
2005	751 154	732 119	714 262	682 851									
2006	757 318	738 847	706 355										
2007	772 909	738 919											
2008	826 885												
2009													
2010													
2011													
2012													
2013													
2014													
2015													
2016													
2017													
2018													
2019													
2020													
2021													

Le triangle des règlements cumulés (k€) « as-if » 2021 de la LoB income protection est donné ci-dessous :

AY\DY	1	2	3	4	5	6	7	8	9	10	11	12	13
1996	186 009	260 210	246 674	239 852	234 297	227 887	222 477	215 710	209 507	205 912	202 463	199 584	194 069
1997	185 144	261 151	252 072	243 745	234 236	228 592	223 540	218 634	214 318	211 281	207 871	202 603	203 273
1998	188 488	272 441	263 408	252 747	245 692	240 646	234 472	229 985	225 451	222 067	214 202	213 749	209 713
1999	176 198	259 007	253 976	241 666	235 697	230 135	224 391	220 814	218 029	211 692	211 197	207 266	202 771
2000	157 140	272 729	264 179	257 096	249 246	242 930	236 806	233 587	226 581	225 214	221 186	216 298	211 321
2001	196 984	276 204	267 604	259 153	257 215	250 672	247 506	240 290	239 166	233 989	227 735	222 217	220 025
2002	188 785	251 032	241 707	235 719	229 681	226 370	219 134	217 530	213 934	208 962	204 567	202 949	201 795
2003	175 457	226 424	222 117	213 812	210 707	203 377	202 277	199 034	194 139	190 359	189 136	188 140	188 236
2004	171 889	228 780	220 495	216 331	208 410	206 921	202 181	196 414	192 685	191 016	189 873	189 780	189 255
2005	174 644	234 781	229 456	221 067	218 226	213 493	208 111	204 094	202 379	201 440	201 233	200 729	198 149
2006	159 020	240 099	228 530	224 628	220 030	212 911	209 679	208 259	207 155	206 714	205 999	203 374	199 451
2007	188 413	232 275	228 965	221 643	215 123	207 982	205 544	204 528	204 158	203 969	200 453	196 876	194 916
2008	175 810	236 296	234 138	225 264	219 874	211 897	210 318	210 284	209 659	207 532	205 299	202 379	201 768
2009	176 879	238 626	238 525	230 418	224 818	224 421	224 384	223 113	220 798	216 201	213 755	212 968	210 040
2010	181 284	252 441	248 465	244 622	241 305	239 318	238 094	234 489	229 079	227 085	226 251	222 692	
2011	185 437	252 657	252 954	252 277	250 067	246 939	243 821	238 915	234 551	233 069	229 043		
2012	182 941	258 784	260 286	256 401	255 219	249 864	242 216	239 333	238 463	235 546			
2013	177 616	262 452	271 775	268 127	263 524	254 103	250 947	248 760	244 235				
2014	189 669	267 957	270 593	266 578	254 477	251 094	247 226	243 504					
2015	176 091	280 110	280 181	269 600	260 924	258 400	254 130						
2016	184 331	275 195	275 613	268 929	265 763	258 399							
2017	215 343	308 256	303 280	300 106	294 886								
2018	236 340	318 495	323 379	324 546									
2019	235 784	347 603	353 772										
2020	241 365	350 491											
2021	263 668												

AY\DY	14	15	16	17	18	19	20	21	22	23	24	25	26
1996	194 067	191 240	186 859	183 246	181 521	180 688	180 655	180 355	178 372	174 890	173 027	172 454	169 645
1997	200 506	196 079	192 030	190 390	189 644	189 959	189 697	187 740	184 025	182 158	181 463	178 554	
1998	205 133	200 545	198 955	198 101	197 918	197 385	195 040	191 728	189 546	188 658	185 652		
1999	198 867	197 146	196 159	195 823	195 370	193 284	189 808	187 637	186 767	184 030			
2000	209 123	207 557	207 668	207 363	205 192	201 436	199 102	198 185	195 064				
2001	218 588	218 385	217 727	214 888	211 081	208 778	207 844	204 676					
2002	201 642	201 111	199 076	195 483	193 289	192 424	189 392						
2003	187 836	186 005	182 466	180 621	179 852	176 947							
2004	187 551	184 040	182 081	181 237	178 400								
2005	194 510	192 321	191 464	188 471									
2006	197 457	196 677	193 837										
2007	194 191	190 932											
2008	198 449												
2009													
2010													
2011													
2012													
2013													
2014													
2015													
2016													
2017													
2018													
2019													
2020													
2021													

Annexe 5 : Méthode de séparation de Taylor

La méthode de séparation de Taylor est une méthode de provisionnement déterministe supposant que les paiements incrémentaux sont déterminés à l'aide d'un paramètre de cadence de règlements et d'un paramètre à effet calendaire :

$$Y_{i,j} = r_j \times \mu_{i+j}$$

Où :

- $Y_{i,j}$: incrément d'année de survenance i et d'année de développement j
- r_j : proportion réglée lors de la $j^{\text{ème}}$ année de développement
- μ_{i+j} : coût total payé lors de l'année calendaire $i + j$

Le triangle incrémental se présente alors sous la forme suivante :

	1	...	j	...	N
1	$r_1 \times \mu_1$...	$r_j \times \mu_j$...	$r_N \times \mu_N$
⋮	⋮			⋮	
i	$r_1 \times \mu_i$		$r_j \times \mu_N$		
⋮	⋮	⋮			
N	$r_1 \times \mu_N$				

Les paramètres $(r_j)_{j \in \{1, \dots, N\}}$ représentent les cadences de paiements des sinistres. Par conséquent, l'estimation des paramètres du modèle est effectuée sous la condition $\sum_{j=1}^N r_j = 1$.

Pour l'année de développement N , nous avons :

$$\mu_N = \sum_{k=1}^N Y_{k, N-k+1}, \quad r_N = \frac{Y_{1,N}}{\mu_N}$$

Pour les années de développement antérieures j :

$$r_{N-j} = \frac{\gamma_{N-j}}{\sum_{k=N-j}^N \mu_k}, \quad \mu_{N-j} = \frac{d_{N-j}}{1 - \sum_{k=N-j}^N r_k}$$

Où :

- $\gamma_{N-j} = \sum_{k=1}^{N-j} Y_{k, N-j+1}$: somme des éléments de la $(N - j) - \text{ème}$ colonne.
- $d_{N-j} = \sum_{k=1}^{N-j} Y_{k, N-k-j+1}$: somme des éléments de la $(N - j) - \text{ème}$ diagonale.

En supposant que le paramètre d'effet calendaire dépend de l'inflation, cette dernière est prise en compte et devient alors un facteur endogène au triangle. La projection du triangle supérieur repose sur la modélisation de l'inflation future.

Annexe 6 : Facteurs de développements calculés par les différents modèles pour les deux LoB étudiées

facteurs de développement - medical expenses				
DY	ResQ	GLM Poisson	GLM Tweedie	as-if 2021
1	1,2391	1,2391	1,2389	1,1982
2	1,0033	1,0027	1,0026	0,9691
3	1,0005	1,0003	1,0003	0,9672
4	1,0001	1,0001	1,0001	0,9668
5	1,0000	1,0000	1,0000	0,9677
6	1,0000	1,0000	1,0000	0,9686
7	1,0000	1,0000	1,0000	0,9702
8	1,0000	1,0000	1,0000	0,9717
9	1,0000	1,0000	1,0000	0,9725
10	1,0000	1,0000	1,0000	0,9730
11	1,0000	1,0000	1,0000	0,9732
12	1,0000	1,0000	1,0000	0,9740
13	1,0000	1,0000	1,0000	0,9745
14	1,0000	1,0000	1,0000	0,9753
15	1,0000	1,0000	1,0000	0,9754
16	1,0000	1,0000	1,0000	0,9755
17	1,0000	1,0000	1,0000	0,9753
18	1,0000	1,0000	1,0000	0,9749
19	1,0000	1,0000	1,0000	0,9748
20	1,0000	1,0000	1,0000	0,9741
21	1,0000	1,0000	1,0000	0,9725
22	1,0000	1,0000	1,0000	0,9711
23	1,0000	1,0000	1,0000	0,9687
24	1,0000	1,0000	1,0000	0,9658
25	1,0000	1,0000	1,0000	0,9560

facteurs de développement - income protection				
DY	ResQ	GLM Poisson	GLM Tweedie	as-if 2021
1	2,2739	2,2751	2,2706	1,4159
2	1,1992	1,1980	1,1947	0,9870
3	1,0955	1,0948	1,0936	0,9755
4	1,0393	1,0401	1,0399	0,9778
5	1,0177	1,0177	1,0176	0,9783
6	1,0104	1,0104	1,0103	0,9826
7	1,0054	1,0054	1,0054	0,9847
8	1,0044	1,0045	1,0045	0,9851
9	1,0024	1,0024	1,0025	0,9870
10	1,0021	1,0021	1,0021	0,9864
11	1,0017	1,0017	1,0018	0,9873
12	1,0015	1,0015	1,0015	0,9881
13	1,0014	1,0014	1,0014	0,9897
14	1,0010	1,0010	1,0010	0,9885
15	1,0012	1,0012	1,0012	0,9895
16	1,0008	1,0008	1,0008	0,9903
17	1,0008	1,0008	1,0008	0,9915
18	1,0008	1,0008	1,0008	0,9917
19	1,0007	1,0007	1,0007	0,9909
20	1,0010	1,0010	1,0010	0,9898
21	1,0007	1,0007	1,0007	0,9874
22	1,0007	1,0007	1,0007	0,9878
23	1,0007	1,0007	1,0007	0,9898
24	1,0010	1,0010	1,0010	0,9902
25	1,0002	1,0002	1,0003	0,9837

Chaque colonne des tableaux ci-dessus représentent une méthode de calcul appliquées :

- ResQ : facteurs de développement issus de l'outil de provisionnement ResQ
- GLM Poisson : facteurs de développement issus de l'application du GLM Poisson après estimation du triangle supérieur
- GLM Tweedie : facteurs de développement issus de l'application du GLM Tweedie après estimation du triangle supérieur
- As-if 2021 : facteurs de développement issus de l'utilisation des triangles « as-if » 2021

TABLE DES FIGURES

Figure 1 – Piliers de la directive Solvabilité 2	19
Figure 2 – Bilan en vision économique	20
Figure 3 – Pieuvre des modules de la directive Solvabilité 2	22
Figure 4 – Modules de risques couverts par le modèle interne de Groupama	23
Figure 5 – Lignes d'activité prises en compte au sein du modèle interne	30
Figure 6 – Calcul du Claim Development Result	33
Figure 7 – Etapes de la méthode Bootstrap	37
Figure 8 – Caractéristiques des GLM étudiés	55
Figure 9 – Critère de durée (en année)	60
Figure 10 – Critère de développement sur les paiements (medical expenses)	61
Figure 11 – Critère de développement sur les charges (income protection)	61
Figure 12 – Critère de développement sur les paiements (income protection)	61
Figure 13 – Résultats du test d'hypothèse d'indépendance	62
Figure 14 – Résultats du test graphique de l'hypothèse de linéarité des 4 premiers couples (medical expenses)	63
Figure 15 – Facteurs de développement et pentes de régression des 4 premiers délais de développement (medical expenses)	63
Figure 16 – Résultats du test graphique de l'hypothèse de linéarité des 6 premiers couples (income protection)	64
Figure 17 – Facteurs de développement et pentes de régression des 4 premiers délais de développement (income protection)	65
Figure 18 – Résultats du test statistique de l'hypothèse de variance (medical expenses)	66
Figure 19 – Résultats du test graphique de l'hypothèse de variance des quatre premières années de développement (medical expenses)	66
Figure 20 – Résultats du test statistique de l'hypothèse de variance (income protection)	67
Figure 21 – Résultats du test graphique de l'hypothèse de variance des quatre premières années de développement (income protection)	67
Figure 54 – Taux d'actualisation en base 2021	69
Figure 53 – Indices du CPI et de l'ONDAM	69
Figure 55 – Résultats du test d'hypothèse d'indépendance entre années de survenance sur les triangles « as-if » 2021	69
Figure 22 – Etape de la procédure Bootstrap (étape 5 modifiée)	71
Figure 23 – Description des impacts de corrélation étudiés	72
Figure 24 – Comparaison des MSEP à un an (en M€)	73
Figure 25 – Comparaison des corrélations implicites entre MSEP	75
Figure 26 – Comportement de la corrélation lors l'impact à priori	76
Figure 27 – Etapes 4 à 6 de la méthode du Bootstrap à un an	76

Figure 28 – Comparaison des corrélations implicites (25%).....	77
Figure 29 – Comparaison des volatilités à un an en fonction des pseudo-facteurs de développement utilisés.....	77
Figure 30 – Matrice de corrélations des rangs entre boni-mali (medical expenses).....	79
Figure 31 – Matrice de corrélations des rangs entre boni-mali (income protection).....	80
Figure 32 – Estimation des paramètres du GLM Poisson (medical expenses).....	82
Figure 33 – Estimation des paramètres du GLM Poisson (income protection).....	83
Figure 34 – Vraisemblance en fonction de la valeur du paramètre p (medical expenses).....	83
Figure 35 – Vraisemblance en fonction de la valeur du paramètre p (income protection).....	83
Figure 36 – Paramètres de puissance de variance.....	84
Figure 37 – Estimation des paramètres du GLM Tweedie (medical expenses).....	84
Figure 38 – Estimation des paramètres du GLM Tweedie (income protection).....	85
Figure 39 – Provisions BE et écart-type à un an.....	85
Figure 40 – Variation de provisions BE et écart-type à un an par rapport aux résultats du Bootstrap à un an.....	87
Figure 41 – Distributions des CDR (medical expenses).....	88
Figure 42 – Minimum, maximum, étendue et moyenne des CDR estimés (medical expenses).....	88
Figure 43 – Distributions des CDR (income protection).....	89
Figure 44 – Minimum, maximum, étendue et moyenne des CDR estimés (income protection).....	90
Figure 45 – Indicateurs de qualité des modèles (medical expenses).....	91
Figure 46 – Indicateurs de qualité des modèles (income protection).....	91
Figure 48 – Courbe de Lorenz.....	92
Figure 49 – Courbes de Lorenz des réserves en fonction des années de survenance (medical expenses).....	93
Figure 50 – Indices de Gini des réserves (medical expenses).....	93
Figure 51 – Courbes de Lorenz des réserves en fonction des années de survenance (income protection)	94
Figure 52 – Indices de Gini des réserves (income protection).....	94

BIBLIOGRAPHIE

- [1] GROUPAMA, *Politique du modèle interne partiel non-vie*
- [2] GROUPAMA, *Corpus méthodologique : calcul du SCR de souscription non-vie risque de réserves et de prime*
- [3] GROUPAMA, *Annexe du corpus méthodologique : calcul du SCR de souscription non-vie risque de réserves et de prime*
- [4] *Cours universitaire*, Formation DUAS (Diplôme Universitaire d'Actuariat de Strasbourg)
- [5] Michael MERZ and Mario V. WÜTHRICH, *Modelling the claims development result for solvency purposes*, Casualty Actuarial Society E-Forum, Fall, 2008
- [6] Alexandre BOUMEZOUED, Yoboua ANGOUA, Laurent DEVINEAU, Jean-Philippe BOISSEAU, *One-year reserve risk including a tail factor: closed formula and bootstrap approaches*, 2011
- [7] Mario V. WÜTHRICH, *Claims reserving using Tweedie's compound model*, 2003
- [8] Arthur CHARPENTIER, Christophe DUTANG, *L'actuariat avec R*, 2012
- [9] Arthur CHARPENTIER, *Actuariat de l'assurance non-vie #10*, 2017
- [10] Arthur CHARPENTIER, *Actuariat de l'assurance non-vie #11*, 2017
- [11] Talel HAMMAMI, *Sur la calibration des corrélations dans un modèle interne non-vie sous Solvabilité 2*, Mémoire d'Actuariat, 2017
- [12] Etienne BUSSON, *Evaluation du risque provisionnement à 1 an : Adaptation de la méthode de Merz & Wuthrich à des cas non standards*, Mémoire d'Actuariat, 2012
- [13] Thomas MACK, *Measuring the Variability of Chain Ladder Reserve Estimates*, 1993
- [14] <http://www.jybaudot.fr/Stats/indicecartes.html> (les indicateurs d'écarts)