

Mémoire présenté devant le Centre d'Etudes Actuarielles
pour l'obtention du diplôme
du Centre d'Etudes Actuarielles
et l'admission à l'Institut des Actuaire
le : 23/06/2011

Par : MR BEN MOHAMED SALAH
Titre : RESULTAT DE COUVERTURE EN ASSURANCE VIE
RISQUES OPERATIONNELS LIES AUX RETARDS DE SAISIE

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

Membre présent du jury de
l'Institut des Actuaire :

signature

Entreprise : AXA FRANCE

Nom : BEN MOHAMED

Signature : 

Directeur de mémoire en entreprise :

Nom : _____

Signature : 

Invité :

Nom : _____

Signature : _____

Membres présents du jury du
Centre d'Etudes Actuarielles :

Thomas BEHAR 

Vincent DAMAS

Gérard CROSET 

Arnaud COHEN

Jean-Pierre DIAZ 

Brigitte DUBUS

Paul ESMEIN

Michel FROMENTEAU

Benoît HUGONIN

Christophe IZART

Pierre PETAUTON

Florence PICARD 

Christian-Yann ROBERT

Secrétariat :

Bibliothèque :

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

Signature du responsable entreprise



Signature du candidat



Salah Ben Mohamed (Promotion 2002)

Résumé

Les règles de calcul des dates de valeur, écrites dans les conditions générales des contrats d'assurance vie, engagent juridiquement l'assureur. Celui-ci obtient l'exécution des ordres financiers à des dates potentiellement différentes chez ses dépositaires selon la remontée de ceux-ci au carnet d'ordres.

Les actes de gestion saisis au niveau du service client entraînent donc un risque financier concrétisé au niveau du résultat de couverture de l'assureur.

Ce mémoire s'intéresse à l'étude des risques liés au résultat de couverture en Unités de Comptes et aux impacts des retards de saisie des services clients sur celui-ci.

La baisse des rendements en Euro et la nécessité d'augmenter la rentabilité orientent de plus en plus les assureurs vie vers de la production en Unités de Compte.

Ce risque constitue donc un enjeu majeur pour la profession particulièrement dans le contexte de Solvabilité II.

Nous définissons le résultat de couverture économique en le distinguant du résultat purement comptable. Nous mettons en évidence qu'il peut être calculé sur chaque flux saisi par les services clients, la somme algébrique constituant alors le résultat de couverture annuel sur le portefeuille considéré. Nous mettons en évidence le poids des saisies rétroactives (retards de saisie en gestion) dans celui-ci. Nous expliquons les différentes causes de retard constatées. Nous modélisons ensuite ce retard (loi mélange d'exponentielles) ainsi que l'évolution des cours (RSLN2).

La modélisation des montants a constitué une difficulté théorique : la queue épaisse de la distribution (des faibles montants très nombreux et des grands montants très rares) n'a pas trouvé de loi usuelle permettant une validation statistique, nous avons dû chercher une loi mélange (log normale Pareto) en nous inspirant des travaux de Planchet et Théron sur le risque de modèle du capital économique pour les indicateurs de queue de distribution (Solvency II).

Les méthodes de Monte Carlo nous permettent alors d'établir une distribution des résultats de couverture annuels.

A partir de celle-ci, on s'attache à étudier la queue de distribution.

L'exigence en capital conforme aux standards Solvabilité II est alors calculée pour un Milliard d' Euros de flux au passif.

La typologie des retards de saisie nous permet de segmenter les risques en souscription vie et risque opérationnel.

Conformément à l'esprit de Solvency II, nous effectuons des recommandations pour réduire notre exposition aux risques et concluons sur les perspectives de poursuite de notre étude.

Abstract

The rules for calculating the value dates written in terms of life insurance contracts legally bind the insurer. It obtains the execution dates of the financial potential among its various stakeholders by the rise of these in the order book.

The management acts before the level of customer service thus involves a financial risk embodied in the result of coverage of the insurer.

This thesis focuses on the study of risks to the result of coverage units and accounts for the impacts of delays before services clients on it.

Lower yields in euro and the need to increase the direct cost of increasing the life insurers to the production units of account.

This risk is therefore a major challenge for the profession especially in the context of Solvency II.

We define the result of economic hedges in distinguishing purely accounting income. We show that it can be calculated on each stream captured by customer service. The algebraic sum constituting the result will cover the annual portfolio. We highlight the weight of seizures retroactive (late seizure management) in it. We explain the different causes of delay identified. Then we model the delay (mixture of exponential law) and the development courses (RSLN2).

Modeling amounts was a theoretical difficulty: the heavy tail of the distribution (many small amounts and a very few number of large amounts) found does not allow a statistical validated usual law, we had to find a law mixture (Pareto lognormal) taking inspiration from the works of Planchet and Theron on risk indicators I (Solvency II). The Monte Carlo methods allow us then to establish a distribution of results of annual coverage. From this, it seeks to study the tail distribution.

The capital requirement standards compliant solvency II is calculated for a billion Euros flow to liabilities, revenues and benefits.

The types of delays before we can segment the life underwriting risk and operational risk.

In the spirit of Solvency II, we make recommendations to reduce our exposure to risk and conclude on the prospects for continuation of our study.

Sommaire

| | | |
|-------------|--|-----------|
| I. | Introduction générale..... | 7 |
| A. | Principe de l'investissement actif-passif en Unités de Compte | 11 |
| 1. | Quelques éléments de finance comportementale..... | 11 |
| 2. | Dates de valeur actif-passif..... | 11 |
| a. | Date de valeur passif | 11 |
| b. | Date de valeur actif | 13 |
| 3. | Processus de passage des ordres dans le carnet..... | 13 |
| a. | Agrégation des flux..... | 13 |
| b. | Dépositaire et passage d'ordres | 14 |
| B. | Méthodologie de l'ajustement actif-passif..... | 14 |
| 1. | Validation des PM (passif)..... | 15 |
| 2. | Relevés de portefeuille (actif) et rapprochement..... | 15 |
| 3. | Analyse des écarts..... | 15 |
| a. | Les erreurs de saisies d'actif | 15 |
| b. | Cut off..... | 15 |
| c. | Acte de réajustement..... | 16 |
| C. | Résultat de couverture économique et résultat de couverture comptable..... | 17 |
| 1. | Définitions | 17 |
| a. | Risque de placement..... | 17 |
| b. | Principe de congruence..... | 17 |
| c. | Ajustement ACAV..... | 17 |
| d. | Ajustement technique | 17 |
| 2. | Résultat de couverture comptable | 17 |
| 3. | Correspondance entre les visions comptable et calculatoire par flux du RCE | 20 |
| II. | Analyse empirique..... | 24 |
| A. | Constitution du fichier d'analyse..... | 24 |
| B. | Calcul du RCE flux par flux..... | 24 |
| C. | Analyse des résultats 2008 | 26 |
| 1. | Etude des valeurs extrêmes | 28 |
| 2. | Etude du RCE par classe d'actifs | 31 |
| 3. | Distribution RCE sur les entrées | 33 |
| 4. | Distribution RCE sur les sorties | 34 |
| 5. | Distribution des retards de saisie..... | 35 |
| 6. | Impact des actes rétroactifs | 35 |
| III. | Modélisation par le processus RSLN2..... | 39 |
| A. | Présentation du processus | 39 |
| B. | Calcul du maximum de vraisemblance | 40 |
| C. | Comparaison avec d'autres modèles | 41 |
| 1. | Le test du rapport de vraisemblance (LRT)..... | 42 |
| 2. | Le critère d'information d'Akaike (AIC) | 42 |
| 3. | Le critère de Shwartz-Bayes (SBC)..... | 42 |
| 4. | Comparaison entre quelques modèles en utilisant les critères de sélection précédents | 42 |

| | | |
|------------|---|-----------|
| 5. | Fonction de répartition globale du modèle RSLN2..... | 43 |
| 6. | Le test Khi-Deux de Pearson :..... | 43 |
| 7. | Le test de Kolmogorov-Smirnov | 44 |
| IV. | ETUDE DE LA STATISTIQUE DES RETARDS..... | 45 |
| A. | Loi exponentielle..... | 46 |
| B. | Loi de Weibull..... | 47 |
| C. | Loi Gamma | 49 |
| D. | Mélange de deux exponentielles | 50 |
| E. | Mélange de trois exponentielles | 50 |
| F. | Justification empirique du modèle retenu | 52 |
| V. | MODELISATION DES RCE 2008 & 2007 | 54 |
| A. | Méthodologie utilisée..... | 54 |
| B. | Rappel des formules de calcul..... | 54 |
| C. | Formule du RCE pour les actes de gestion rétroactifs | 55 |
| D. | L'approche monte Carlo | 55 |
| 1. | Comment simuler un retard et un scénario financier ?..... | 56 |
| 2. | Structure des flux 2008 | 57 |
| 3. | Distribution du RCE 2008 | 58 |
| a. | Estimation des fréquences de saisie | 58 |
| b. | Distribution par tranche de RCE observés..... | 60 |
| 4. | Etude de la distribution réalisée à partir des cash flows 2007 | 63 |
| a. | Structure des flux 2007 | 63 |
| b. | Distribution par tranche de RCE observés..... | 63 |
| 5. | Comparaison des distributions 2007 et 2008 standardisées | 66 |
| a. | Standardisation..... | 66 |
| b. | Distributions 2007 & 2008 standardisées | 66 |
| c. | Le test de Wilcoxon | 68 |
| d. | Le test de Kruskal-Wallis | 69 |
| VI. | MODELISATION DES FLUX FINANCIERS DES ASSURES..... | 69 |
| A. | Fonction de répartition empirique..... | 70 |
| B. | Lois usuelles | 71 |
| 1. | Loi exponentielle..... | 71 |
| 2. | Loi gamma | 71 |
| 3. | Loi log normale..... | 72 |
| 4. | Loi de Burr de type XII | 72 |
| 5. | Loi de Pareto..... | 72 |
| a. | Avec le minimum observé..... | 72 |
| b. | Avec les méthodes d'optimisation | 73 |
| c. | Régression linéaire de la fonction de survie..... | 73 |
| C. | Le modèle mélange Log normal Pareto..... | 74 |
| 1. | Fonction de survie | 74 |
| 2. | Estimation des paramètres du modèle | 75 |
| 3. | Simulation de la loi mélangée | 76 |
| 4. | Application numérique | 77 |
| 5. | Test d'adéquation | 77 |
| D. | Nouvelle simulation de la distribution des résultats de couverture..... | 77 |

| | |
|---|-----------|
| VII. RISK MANAGEMENT ET RESULTAT DE COUVERTURE | 80 |
| A. Solvabilité I | 80 |
| B. Solvabilité II (Solvency II)..... | 81 |
| 1. Calcul du SCR et value at risk sur la distribution 2008 standardisée à un milliard de flux | 81 |
| 2. Risk management et CTE..... | 82 |
| 3. Réévaluation du modèle avec la loi des retards de longue durée..... | 84 |
| 4. Réévaluation du modèle avec la loi des retards de courte durée | 85 |
| 5. Réévaluation du modèle avec la loi des retards de moyenne durée..... | 86 |
| Conclusion..... | 88 |
| Bibliographie..... | 90 |
| Annexes | 91 |

Remerciements

Je tiens à remercier tout particulièrement Mohamed BACCOUCHE, Directeur des comptes techniques et financiers vie d'AXA France pour toute l'aide et le support qu'il m'a apporté. Merci à Mr Christian ROBERT de l'institut des actuaires ainsi que Mme VALLAT et Mme Françoise PARIS d' AXA France pour leurs lectures et leurs conseils respectifs.

Je remercie également ma famille pour la patience et la compréhension dont elle a fait preuve.

I. INTRODUCTION GENERALE

Historiquement, les problématiques d'adéquation entre l'actif et le passif des compagnies d'assurances vie ne concernaient que les fonds en Euros. En effet, l'objectif principal consistait à ce que les investissements de l'assureur soient capables de servir dans la durée les rendements conformes à sa politique de taux. Dans un contexte général de baisse des rendements sur les fonds en euros, la part des supports en UC devient de plus en plus importante. Les contrats d'assurance multi-supports se sont peu à peu imposés avec une offre de fonds en unités de compte. Ils étaient à l'origine essentiellement distribués par les réseaux captifs (propres à l'assureur), le nombre de supports était limité, ces derniers étaient le plus souvent gérés directement par l'assureur via sa filiale de société de gestion. L'adéquation de l'actif en unités de compte sur le passif était donc une activité relativement fluide et maîtrisée. Le contexte a changé depuis 15 ans : les assureurs vendent de plus en plus d'unités de compte dans le cadre de leurs contrats d'assurance vie.

Ceux-ci offrent un meilleur potentiel de rendement sur le long terme, mais le risque de moins value est entièrement supporté par l'assuré. Ce risque s'apprécie par rapport à une date de valeur de conversion des montants initiaux versés par le client. L'assureur vie s'engage sur la valeur liquidative de cette date, appelée aussi date d'effet, pour convertir en nombre de parts chaque flux d'investissement ou de désinvestissement demandé par l'assuré. Pour ne pas subir à son tour de risque de moins value, l'assureur doit autant que faire se peut s'assurer qu'il possède, à tout moment dans son actif et pour chaque support, les mêmes nombres d'unités de compte que ceux détenus par ses clients au travers de leurs contrats (principe de congruence article A343-1 du code des assurances). Ces opérations sont dites de couverture en UC.

D'autre part, ces opérations doivent être réalisées idéalement de façon simultanée afin de concerner les mêmes valeurs liquidatives. Les dépositaires avec lesquels travaillent les assureurs exigent en général exclusivement des ordres en parts et les dates de valeur sont en général postérieures ou égales au jour de passation des ordres.

Par ailleurs, pour les OPCVM qui constituent la majorité des fonds proposés, la valeur liquidative d'un jour n'est connue qu'au mieux deux jours ouvrés plus tard, les gestionnaires de fonds évaluant les différentes composantes de leurs placements et prélevant différents frais. L'assureur ne connaît pas la valeur liquidative dont il a besoin pour chaque support et estime donc le nombre de parts qu'il vendra ou achètera cependant à la même date que son client. A posteriori, il corrige son estimation par un flux de régulation et ceci se solde par un résultat de couverture associé au flux.

Le résultat de couverture de l'assureur enregistré en résultat technique représente, de façon agrégée pour l'ensemble du portefeuille, le coût ou le gain issu du solde de l'ensemble des opérations de couverture des engagements en UC de l'assureur.

Dans un contexte de plus en plus concurrentiel, l'assurance vie s'est élargie aux réseaux non propriétaires et à des supports de plus en plus complexes. En effet, se sont ouverts à la commercialisation des titres vifs (action, obligation ou titre à cotation continue), des fonds structurés, des EMTN (Euro Medium Term Note) gérés par des acteurs externes à la compagnie d'assurances ayant leur propre exigence de fonctionnement. A titre d'exemple, sur les réseaux non propriétaires d'AXA France, plus de 3000 supports sont disponibles issus de plus de 250 sociétés de gestion.

Par ailleurs, la gestion de tous ces supports génère des dysfonctionnements, les circuits d'information étant démultipliés et différenciés (problématiques OST, fusion absorption de support, coupons à distribuer), problématiques de rattachement (paramétrage des supports dans les systèmes d'information). Le non fonctionnement de ces process au passif peut entraîner un déphasage substantiel entre l'exécution de ces ordres au passif et leur réalisation à l'actif, générant un risque financier pour l'assureur.

Par ailleurs, des erreurs humaines de gestion, des dérogations commerciales sur les dates de valeur, des retards de traitement peuvent perturber le fonctionnement normal de l'ajustement du passif en UC sur l'actif à conséquences financières d'autant plus fortes selon le cas concerné (retard important, flux financier conséquent, fluctuation des cours...).

Cela se mesure par le nombre d'actes rétroactifs réalisés au passif.

A cela vient se greffer des dysfonctionnements des systèmes d'information, entraînant des actes non réalisés ou effectués avec retard au passif, créant là encore un déphasage temporel pouvant générer un résultat de couverture non négligeable.

Nous allons réaliser une modélisation statistique de ces phénomènes et de leurs impacts financiers. Nous mettrons en évidence que les exigences en capital se montent à plusieurs millions d'euros (32 M€ minimum pour un milliard de flux en montant)

La maîtrise de ce résultat est un enjeu majeur pour les compagnies d'assurances.

D'une part, le cadre européen de Solvency II leur impose de maîtriser au plus près leurs risques techniques (dont le résultat de couverture) et d'autre part, la pression concurrentielle et les contextes de marché les poussent à générer du chiffre d'affaires en UC de plus en plus élevé, ce qui amène ce poste à prendre une importance croissante.

Nous rappellerons dans une première partie l'ensemble de ces définitions et processus de fonctionnement de la couverture dans le groupe AXA en particulier pour un portefeuille d'un réseau non propriétaire; mais ces pratiques sont généralisables à la plupart des assureurs. (I.a & I.b). Nous y expliciterons les différents éléments qui sont à l'origine du résultat de couverture comptable. Nous distinguerons le résultat de couverture qui a un sens économique réel pour l'assureur du résultat purement comptable dû à des problèmes d'évaluation des ajustements techniques (passif) ou Acav (actif). (I.c).

Nous démontrerons, qu'abstraction faite du résultat purement comptable, il correspond à la somme algébrique des résultats de couverture calculés sur chaque flux. . (I.c).

Dans une deuxième partie, nous analyserons le résultat de couverture annuel 2008 après avoir mis en évidence les différents éléments à prendre en compte dans le calcul du résultat sur un flux (cash flow, valeurs liquidatives, écart entre date de saisie et date de valeur).

En effet, le risque peut être aggravé si les actes de gestion sont saisis à effet rétroactif dans le service client (date de valeurs différentes au passif et à l'actif). Nous démontrerons le poids des actes saisis à effet rétroactif dans le résultat de couverture économique annuel 2008.

Nous établirons alors une cartographie des retards de saisie et des explications associées sur 2008. (II)

A partir de la IIIe partie, nous chercherons à modéliser ce risque en appliquant les méthodes de Monte Carlo. Après avoir établi un modèle d'évolution des valeurs liquidatives (II), nous rechercherons une loi adaptée à l'historique de nos retards (III).

Nous générerons alors plusieurs milliers de scénarios nous permettant d'établir la loi des résultats de couverture après avoir constaté la convergence.

Nous réitérerons le même exercice avec les cash flows 2007.

Nous constaterons qu'en calquant la part des cash flows positifs sur celle de 2008, nous retrouvons la même distribution translatée. (IV).

Nous mettrons alors en évidence que la distribution des montants peut être modélisée par une loi log normale Pareto. (V)

Nous nous attacherons alors à calculer des indicateurs de queue de distribution (Value at risk et CTE) et à qualifier les risques encourus. (VI)

Nous réitérerons alors nos simulations Monte Carlo sur la base de plusieurs échantillons annuels de notre loi des montants.

Enfin, nous chercherons à savoir comment varie la distribution des résultats de couverture en fonction de différents paramètres de la loi de retard.

Nous mettrons en évidence les différents niveaux de capital exigible selon la nomenclature Solvency II en distinguant les risques mis en évidence par nature (risque de souscription & risque opérationnel).

1^{ère} Partie

Principes de l'adossement actif/passif
en Unités de compte et des résultats de couverture

A. Principe de l'investissement actif-passif en Unités de Compte

1. Quelques éléments de finance comportementale

C'est indéniablement la recherche de la performance qui va pousser nos assurés à investir sur des unités de compte. Cela peut provenir d'une espérance de gain liée à la spéculation sur la valeur de l'actif lui-même ou tout simplement un coupon ou dividende dont l'assuré attend (fonds obligataires) ou espère une meilleure rémunération que le taux de rendement (participation aux bénéfices comprise) des fonds en Euros. En contrepartie, l'assuré assume un certain risque et va chercher à le minimiser. Il va en général suivre les recommandations de son conseiller en fonction de son aversion aux risques, de ses objectifs de rendement, de son horizon de placement. Il va donc choisir d'investir sur les supports dont il anticipe la hausse ou la stabilisation des cours (à rendement supérieur aux fonds Euros). De façon symétrique, il va vendre les unités de compte dont il anticipe la baisse de la valeur. Ces anticipations et réactions de l'investisseur sont appréhendées par les théories de la finance comportementale :

- Les investisseurs ont tendance à aller sur les unités de compte lorsque leurs cours montent. En d'autres termes, les bénéfices présents sont dans la psychologie de l'investisseur un présage favorable de rendements futurs. C'est la théorie de l'argent du casino (« house money effect »).
- L'aversion pour les pertes est en général plus grande que l'appétit pour les gains (théorie des perspectives de Kahneman et Tversky) : l'investisseur va choisir le risque d'une perte future potentielle plus élevée plutôt qu'une moins-value moins importante mais certaine aujourd'hui. A l'opposé, il choisira un gain certain moins important qu'une espérance de gain plus élevé plus tard.
- Les versements programmés ont pour but de sécuriser l'investissement global en lissant le cours d'achat des Unités de Compte à la baisse. Certaines études tendent à montrer que l'espérance de gain selon les scénarios de cours est assez proche lorsqu'on compare un investissement unique et le même montant réparti sur plusieurs versements programmés. Ces derniers minimisent le risque de perte à espérance de gain voisine. Par ailleurs, ils permettent de décharger l'assuré de l'initiative de l'opération et lui donnent une discipline de versements qui s'adapte à sa capacité d'épargne (mensuelle, trimestrielle....).
- On observe assez régulièrement un certain mimétisme dans le comportement des investisseurs.

2. Dates de valeur actif-passif

a. Date de valeur passif

Lors de la signature d'un contrat d'assurance vie, un exemplaire des Conditions Générales est remis au client. En général, dans ce dernier un ou plusieurs articles sont consacrés à la définition des dates de valeur passif : ce sont celles dont les valeurs liquidatives correspondantes servent à convertir en nombre de parts le montant investi ou désinvesti par le client sur la provision mathématique du support choisi.

Cette date engage l'assureur. Quelle que soit la date à laquelle il aura effectivement réalisé l'opération pour le compte de son assuré au niveau de son actif en titres, c'est bien cette valeur liquidative qui fera foi pour la conversion du flux au niveau de l'épargne du client.

Dates de valeur effectives pratiquées dans les conditions générales

Les dates de valeur retenues pour les produits phares de notre étude diffèrent en fonction de l'acte et du moyen de paiement du souscripteur.

Ci-dessous un extrait des conditions générales concernant les dates de valeur :

Dates de valeur appliquées à chaque opération

L'investissement du versement initial ne peut être effectué qu'après réception par le Service Clients de l'assureur du dossier complet avec notamment le Bulletin de Souscription dûment rempli et signé.

Tout investissement ou désinvestissement ne peut être effectué sur les supports en Unités de Compte que sur la base d'un cours ou d'une valeur de l'unité de compte inconnu, c'est à dire déterminé après réception de la demande complète correspondante, dans les conditions indiquées ci-dessous.

Pour chacun des événements suivants, la date de valeur* retenue est :

- pour les versements de primes effectués par virement, le 1er ou le 2ème jour ouvré (1) qui suit le jour de réception par le Service Clients de l'assureur de l'avis de virement du versement permettant à l'assureur d'identifier cette prime, et du Bulletin de Souscription (ou de la demande de versement complémentaire),
- pour les versements de primes effectués par chèque, le 3ème ou le 4ème jour ouvré (1) qui suit le jour de réception par le Service Clients de l'assureur du chèque et du Bulletin de Souscription (ou de la demande de versement complémentaire),
- pour le calcul des sommes dues en cas de rachat de l'assuré, le 1er ou le 2ème jour ouvré (1) qui suit le jour de réception par le Service Clients de l'assureur de la demande complète de rachat,
- pour le calcul des sommes dues en cas de décès de l'assuré, le 1er ou le 2ème jour ouvré (1) qui suit le jour de réception par le Service Clients de l'assureur de l'acte de décès original et de la demande complète de règlement pour au moins un bénéficiaire,
- pour les réorientations d'épargne, le 1er ou le 2ème jour ouvré (1) qui suit le jour de réception par le service Clients d'une demande complète de réorientation d'épargne. Les opérations d'investissement ou désinvestissement des Unités de Compte concernées par la réorientation de l'épargne sont simultanées, si leur rythme de cotation ou valorisation le

permet et si les fonds sont disponibles sans délai pour l'assureur. Dans le cas contraire, l'investissement dans un support en Unités de Compte serait réalisé le 1er jour, ou au plus tard le 2ème jour ouvré qui suit l'opération correspondante de désinvestissement.

Cependant si l'assureur se trouvait dans l'impossibilité d'acheter ou de vendre des Unités de Compte dans les conditions ci-dessus, seront utilisées les valeurs auxquelles l'assureur aura pu acheter ou vendre celles-ci.

Les situations de contrat demandées en cours d'année sont établies sur la base des dernières valeurs connues des Unités de Compte à la date de la demande.

(1) « Jour ouvré » désigne un jour de cotation ou de valorisation (ces jours devant être des jours ouvrés pour l'assureur).

On peut remarquer que :

- La date de valeur est liée à la complétude du dossier, ce qui suppose que l'acte ne sera saisi que lorsque l'ensemble des pièces nécessaires sera fourni. Un certain nombre d'allers retours peuvent donc avoir lieu entre le service client et le souscripteur ou le commercial, l'acte de gestion faisant alors l'objet d'une « attente ». On verra par la suite que cela aura un impact sur les retards d'investissement entre l'actif et le passif.
- D'autre part, nous voyons que la date de valeur est liée au moyen de paiement, lorsque le paiement est effectué par chèque, un certain temps est nécessaire pour encaisser les fonds, ceci se traduit par une date de valeur de trois ou quatre jours postérieure à la date de réception du dossier au service client. La plupart des carnets d'ordres savent gérer cette dichotomie en faisant remonter les ordres après l'encaissement effectif des versements effectués par le client.

b. Date de valeur actif

C'est la date de valeur liquidative retenue par les dépositaires pour réaliser les ordres de vente ou d'achat des unités de compte correspondants à la contrepartie des flux demandés par les clients sur leurs contrats.

En général, cette date correspond au jour de réception des ordres par le dépositaire agréé d'un horaire maximal de réception.

Sur le portefeuille étudié, nous avons plus de cinquante dépositaires.

Le risque financier existe à partir du moment où les mêmes montants ne sont pas convertis à la même valeur liquidative au passif et à l'actif.

3. Processus de passage des ordres dans le carnet

a. Agrégation des flux

Le carnet d'ordre est un applicatif du système de gestion dont la raison d'être est de traduire en ordres financiers les flux saisis sur les supports en UC des contrats de nos clients. Chaque jour et pour chaque support la somme algébrique des mouvements dont le jour d'effet est le lendemain remontent et donnent lieu à un ordre d'achat ou de vente. Les actes rétroactifs c'est-à-dire à date d'effet antérieure aboutissent immédiatement au carnet d'ordre et sont aussi additionnés à la somme précédente pour chaque support.

b. Dépositaire et passage d'ordres

Les dépositaires passent les ordres en parts.

Lorsque l'ensemble des flux devant être réalisés à une date d'effet donnée sont agrégés, un ordre global remonte au carnet la veille (ou le premier jour ouvré précédent) par dépositaire et support. Pour estimer le nombre de parts concerné par l'ordre (puisque qu'on ne connaît pas la valeur liquidative du lendemain), le carnet d'ordres utilise la dernière Valeur liquidative connue. Ce choix d'implémentation est lié au fait qu'on considère que celle-ci sera la plus proche en moyenne de la valeur liquidative cible, et donc générera l'erreur la plus faible au niveau de l'ordre passé.

Si on analyse la situation du point de vue d'un ordre correspondant à un flux unique au passif (à l'agrégation près),

l'ordre sera donc normalement exécuté à la même date au passif et à l'actif, mais le flux en parts sera différent. Dans le premier cas, le flux sera converti à la valeur liquidative d'effet une fois celle-ci connue; dans le deuxième cas, l'achat ou la vente sera réalisé avec la même valeur liquidative d'effet, mais sur un nombre de parts estimé forcément différent du passif sauf en cas de stabilité de la valeur liquidative.

Cette méthode a l'avantage de limiter les risques financiers inhérents à la réalisation d'opérations financières effectuées avec des cours différents, mais on voit aisément qu'elle a l'inconvénient de pouvoir générer un décalage entre l'actif en UC de l'assureur et sa dette envers ses assurés (passif en Unités de Compte).

Pour cela et selon les carnets d'ordre en vigueur sur le marché, il existe parfois un flux de régulation ou réajustement qui va compléter l'ordre passé à l'actif à hauteur du flux correspondant au passif (dans un sens ou dans l'autre).

Là encore, selon les systèmes d'information, ce flux est soit passé automatiquement presque immédiatement après le flux qu'il corrige avec remontée directe au carnet d'ordres, soit calculé mais ne donnant lieu à passage d'ordres que suite à intervention humaine.

Le portefeuille que nous étudierons se situe dans ce dernier cas.

La correction ne peut alors avoir lieu qu'après constatation du désadossement actif-passif par une équipe ou personne dédiée à l'analyse de cette problématique.

B. Méthodologie de l'ajustement actif-passif

Cette partie a pour but de décrire de façon succincte l'analyse des écarts de désadossement actif passif telle qu'elle est réalisée pour notre portefeuille et les différentes typologies d'écarts constatés et de correction associées.

L'analyse est mensuelle pour tous les portefeuilles d'AXA et pour la majorité des assureurs de la place.

1. Validation des PM (passif)

Le passif en UC pour un support donné correspond aux provisions mathématiques de tous les contrats en portefeuille à la fin du mois d'analyse.

L'analyse d'un mois M a généralement lieu courant M+2 afin de consacrer M+1 à la validation des provisions mathématiques dont on s'assure qu'elles correspondent à nos engagements vis-à-vis de nos clients à chaque fin de mois.

Il est extrait du système de gestion choisi et est segmenté par dépositaire puisque la maille d'analyse est le couple support, dépositaire.

2. Relevés de portefeuille (actif) et rapprochement.

La saisie du nombre de parts à l'Actif est faite manuellement chaque mois.

Pour chaque support, il faudra rechercher dans les relevés titres téléchargés ou papier, le nombre de parts de chaque dépositaire et le renseigner à la même maille qu'au passif : couple support, dépositaire.

Notre portefeuille concerne des réseaux non propriétaires pour lesquels il existe plus de 50 dépositaires différents et environ 3000 supports.

Cette opération est longue et chronophage. Ceci explique aussi pourquoi l'analyse du mois M a lieu courant M+2.

3. Analyse des écarts

L'analyse va consister à chercher, expliquer et corriger les écarts d'encours en parts sur les positions actifs et passifs. On recherche par ordre de priorité :

a. Les erreurs de saisies d'actif

Vérifier qu'il n'y a pas d'erreurs de saisie de l'actif à partir du relevé titres du mois de l'analyse.

b. Cut off

Lors de l'analyse, on peut distinguer deux types de cut-off :

1. Les cut-off qui correspondent aux ordres de fin de mois M mais qui ne seront réellement passés qu'en M+1

Afin d'identifier ce type de cut-off, il faudra reconstituer l'actif du mois M-1 à M.

$$\text{Actif de } M = \text{Actif de } M-1 + \text{ordres d'achat } M - \text{ordres de vente } M$$

Les différents ordres d'achat et de vente s'afficheront et permettront de retrouver l'écart entre les deux périodes.

Si on ne parvient pas à reconstituer l'écart, il faudra consulter les relevés espèces pour identifier les ordres réellement passés. En d'autres termes, il va falloir comparer les nombres de parts d'achat ou vente (ou parfois les montants) des relevés espèces à ceux présents sur le système de gestion.

2. Les cut-off qui correspondent aux ordres de M+1 mais qui ont déjà été pris en compte auparavant au passif.

Ici on parlera de cut-off lorsque les dates de l'actif et du passif concernant un même acte se retrouvent sur deux mois consécutifs.

c. Acte de réajustement

Il existe trois types d'actes de réajustement :

- les frais de gestion,
- les régularisations sur flux.
- Les opérations sur titres.

1. Frais de gestion

Sur l'ensemble des contrats gérés, Axa se rémunère en prélevant des frais de gestion (ou d'encours) de l'ordre d'environ 1% par an.

Ces flux ne donnent pas lieu à des ordres automatiques à l'actif puisqu'ils concernent la relation assureur-assuré. Ils sont calculés contrat par contrat et support par support par les équipes en charge de l'adossement et donnent lieu à des corrections pour réduire les écarts constatés.

2. Régularisation sur flux

Lorsque l'écart est insuffisamment corrigé, on va rechercher les régularisations sur flux évoquées plus haut, qui seront extraites et agrégées algébriquement (pour bénéficier de compensations). Elles seront traduites en ordre de correction à passer après validation du comité actif-passif qui a lieu autour du 25/M+2 pour une analyse à fin M.

3. Opérations sur titres

Ce sont généralement :

des détachements de coupons paramétrés automatiquement au passif, tandis que leur contrepartie à l'actif doit faire l'objet d'une opération manuelle

ou une fusion-absorption d'un support par un autre refusée par la direction technique sur les contrats de nos assurés pour des considérations juridiques ou de risques assurantiels. L'écart se justifie alors sans donner nécessairement lieu à une correction.

C. Résultat de couverture économique et résultat de couverture comptable

1. Définitions

Rappelons quelques éléments préalables qui vont nous aider à comprendre la notion de résultat de couverture.

a. Risque de placement

Dans les contrats en unités de compte, bien que les actifs détenus soient toujours la propriété de l'assureur, le risque est assumé par l'assuré qui accepte de voir le capital souscrit dans son contrat suivre à la hausse comme à la baisse, l'évolution de l'unité de compte retenue.

b. Principe de congruence

La commercialisation de contrats d'assurance en unités de compte suppose qu'il existe une relation étroite entre les engagements souscrits par les assurés et les actifs admis en représentation de ces mêmes engagements.

L'article R 332-5 du code des assurances précise en effet que les provisions mathématiques des contrats d'assurance sur la vie ou de capitalisation à capital variable, dans lesquels la somme assurée est déterminée par rapport à une valeur de référence, doivent être représentées à l'actif du bilan par des placements entrant dans la composition de cette valeur de référence et dans les proportions fixées par la dite composition. Son application a par ailleurs été précisée par les circulaires de la direction des assurances B3 195 & A1 2221.

c. Ajustement ACAV

Il s'agit de la dénomination donnée à l'écart constaté sur la valeur de la part entre deux inventaires, ou un inventaire et la date de sortie du contrat. L'assureur enregistre en comptabilité, dans des comptes particuliers intitulés à cet effet « Ajustements ACAV plus values ou Ajustements ACAV moins values, la somme de tous ces écarts, qu'ils soient positifs (cas des plus values) ou négatifs (cas des moins values). Ces comptes font partie du résultat technique des contrats (et non pas de leur résultat financier comme c'est le cas pour les plus ou moins values classiques de cession enregistrées lors de la réalisation d'un placement financier).

d. Ajustement technique

Le même calcul est réalisé au niveau du portefeuille donnant lieu aux ajustements techniques.

2. Résultat de couverture comptable

Le résultat de couverture comptable se déduit de la comparaison des ajustements techniques déterminés au passif avec les ajustements ACAV élaborés à l'actif.

Avec les notations suivantes :

Ouv = ouverture, clôt = clôture

A = actif, P = passif

PM = Provision Mathématiques (passif)

VB = Valeur Boursière (actif)

Les ajustements techniques représentent la valorisation faite au passif (sur les comptes clients) pendant la période définie :

Ajust Tech = ajust sur flux + ajust sur stock

Avec **ajust sur flux = NbpartsFluxP*(VL_clot - VL_client)**

Et **ajust sur stock = Nbparts_ouvP*(VL_clot - VL_ouv)**

Les ajustements ACAV représentent la valorisation faite à l'actif (sur les positions prises par l'entreprise) pendant la période définie :

**Ajust ACAV = Nbparts_ouvA*(VL_clot - VL_ouv)
+ NbpartsFluxA*(VL_clot - VL_invest)**

Le RCC est le résultat de couverture comptable, il est défini tel que :

RCC = Ajust ACAV - Ajust tech.

C'est le résultat qui est enregistré dans les comptes au final.

Il se décompose en deux termes :

- Le RC purement comptable qui est dû à une différence d'adossement entre clôture et ouverture.
- Le RCE (résultat de couverture économique) qui représente un gain ou une perte sur la trésorerie de l'entreprise dû à l'écart de montant entre les flux à l'actif et les flux au passif.
- Lorsque le RC est nul, il y a identité entre les deux.

Illustrons cela avec un cas d'école :

Supposons que le portefeuille UC de l'assureur se réduise à un support et que les mouvements annuels se cantonnent à un flux d'achat. Lorsque le portefeuille est parfaitement adossé, le résultat de couverture calculé comme un solde entre deux ajustements correspond bien à l'expression d'une perte ou d'un gain économique. L'achat de la demi-part s'étant fait à 13 euros au lieu de 12 au passif, tout se passe comme si la compagnie avait acheté la moitié d'une part à 13 euros et l'avait vendue à 12 au client.

| | Ajustements passif techniques | | |
|----------------------|-------------------------------|-----------------|--------------|
| | valeur liquidative | Nombre de parts | Montant en € |
| ouverture | 10,0 | 1,0 | 10,0 |
| Flux | 12,0 | 0,5 | 6,0 |
| clôture | 14,0 | 1,5 | 21,0 |
| ajustement technique | | | 5,0 |

| | Ajustements actif ACAV | | |
|-----------------|------------------------|-----------------|--------------|
| | valeur liquidative | Nombre de parts | Montant en € |
| ouverture | 10,0 | 1,0 | 10,0 |
| Flux | 13,0 | 0,5 | 6,5 |
| clôture | 14,0 | 1,5 | 21,0 |
| ajustement acav | | | 4,5 |

RCC - 0,50

Lorsque les positions à l'ouverture et/ou à la clôture sont différentes, nous sommes alors dans un cas de sur-adossement ou de sous-adossement selon le sens de l'écart constaté :

le RCC est égal à la somme du RCE et du RC, soit dans l'exemple suivant $-0,5 - 4 = -4,5$.

Le résultat de couverture est le fruit d'une réalité économique et de l'expression d'un désadossement actif-passif constaté qu'il soit réel ou temporaire (produit d'un cut off dû à l'arrêté comptable).

| | valeur liquidative | Nombre de parts | Montant en € |
|----------------------|--------------------|-----------------|--------------|
| ouverture | 10,0 | 2,0 | 20,0 |
| Flux | 12,0 | 0,5 | 6,0 |
| clôture | 14,0 | 2,5 | 35,0 |
| ajustement technique | | | 9,0 |

| | Ajustements actif ACAV | | |
|-----------------|---------------------------|-----------------|---------------|
| | valeur liquidative | Nombre de parts | Montant en € |
| ouverture | 10,0 | 1,0 | 10,0 |
| Flux | 13,0 | 0,5 | 6,5 |
| clôture | 14,0 | 1,5 | 21,0 |
| ajustement ACAV | | | 4,50 |
| RCC | | | - 4,50 |

3. Correspondance entre les visions comptable et calculatoire par flux du RCE

En partant d'une situation parfaitement adossée au départ, nous allons démontrer que l'on peut retrouver le montant du RCE imputable à chaque flux et donc calculer le RCE réel par agrégation de ceux calculés sur les flux, en nous basant sur le fonctionnement du carnet d'ordre.

Prenons pour cela un exemple réel constaté dans notre carnet d'ordre.

Soit une vente de 1990 euros réalisée le 14/03/08 sur un support de notre portefeuille dont la dernière valeur liquidative connue au moment de la saisie (notons-la vlsaisie) était de 91.22 Euros. L'ordre de vente se fonde donc sur un nombre de parts estimé de 21.81 parts.

La valeur liquidative de la date d'effet au passif (notons la vleffet) est de 86.46 euros.

Le nombre de parts réel racheté sur le contrat du client est donc de 23.016.

Lors de l'ajustement de fin mai 2008, le flux d'ajustement pour assurer la congruence actif passif est donc la vente de $23.016-21.81=1.2$ parts.

A la date d'ajustement du 26/05/08, la valeur liquidative du support est de 99.83 €.

Le RCE vaut donc 16.05 €.

En résumé avec les notations suivantes :

$$\text{Regul} = (\text{montant/vlsaisie}) - (\text{montant/vleffet}) = ((-1990/91.22) - (-1990/86.46)) = -1.2$$

$$\text{Rce} = \text{regul} * (\text{vlajust} - \text{vleffet}) = -1.2 * (86.46 - 99.83) = 16.05 \text{ €}.$$

L'assureur gagne donc 16.05 € sur cette opération.

Nous allons maintenant écrire ces opérations en vision compte de résultat comme décrit plus haut.

Supposons pour plus de clarté que l'on réalise les comptes à la date d'ajustement du flux, que ce dernier soit le seul de l'exercice et que la compagnie parte d'une situation avec ce seul support et nous supposerons également que l'encours de départ est égal à 50 parts valorisées à 90 euros (le résultat ne change pas si on modifie l'encours initial, parfaitement adossé).

Il vient alors :

| | Ajustements techniques passif | | |
|----------------------|-------------------------------|-----------------|--------------|
| | valeur liquidative | Nombre de parts | Montant en € |
| ouverture | 90,00 | 50,00 | 4500,00 |
| Flux | 86,46 | -23,02 | - 1990,00 |
| clôture | 99,83 | 26,98 | 2693,77 |
| ajustement technique | | | - 183,77 |

| | Ajustements acav actif | | |
|------------------------|------------------------|-----------------|--------------|
| | valeur liquidative | Nombre de parts | Montant en € |
| ouverture | 90,00 | 50,00 | 4500,00 |
| Flux | 86,46 | -21,82 | - 1886,16 |
| Flux ajustement | 99,83 | - 1,20 | - 119,90 |
| clôture | 99,83 | 26,98 | 2693,77 |
| ajustement ACAV | | | - 199,83 |

RCE 16,06

Nous voyons donc qu'il est possible en théorie de reconstituer le résultat de couverture économique imputable aux actes financiers classiques en le calculant flux par flux et en réalisant la somme algébrique des ces montants.

Approche comparée et limites de l'étude

On ne s'intéressera qu'au résultat de couverture purement économique dans le cadre de cette étude.

Le résultat de couverture purement comptable ne concerne pas les événements ayant pour cause directe le service client.

Nous limitons l'étude au résultat de couverture concerné par les risques imputables au service client et à l'activité de gestion des contrats (humains ou techniques).

2eme Partie

Analyse du résultat de couverture économique

II. ANALYSE EMPIRIQUE

A. Constitution du fichier d'analyse

Pour reconstituer le résultat de couverture économique, nous avons donc décidé de le recalculer flux par flux.

Nous sommes partis des bases de données 2008 que nous avons exploitées sous SAS, étant donné le grand nombre de flux financiers de notre portefeuille d'analyse AXA.

Un flux financier est la somme algébrique des montants demandés par nos assurés au passif par contrat par support (un même code ISIN), par type d'acte, par date d'effet passif et par date de saisie.

En effet, ces éléments caractérisent totalement le résultat de couverture économique puisque les dates actif et d'ajustement se déduisent de la date de saisie.

NB : L'ensemble des mouvements issus des actes de gestion par contrat/support dépasse les 90 000 occurrences en 2008.

Sur 2008, nous n'avons conservé que les flux financiers dont nous avons récupéré l'ensemble des valeurs liquidatives nécessaires au calcul du RCE soit 89 297 flux (la table en contenait 91 538 à l'origine).

Le nombre de supports étudiés est de 985.

Ce portefeuille correspond à un réseau non propriétaire d'AXA. Les produits vendus sur ces réseaux très concurrentiels se doivent d'être le plus attractif possible à l'égard d'un client potentiel. Ceci explique la grande diversité de supports proposés. Tous les plus grands gestionnaires de fonds de la place sont représentés.

La récupération de toutes les valeurs liquidatives nécessaires pour tous les flux et tous les supports a été très chronophage. On a ensuite calculé le RCE attribuable à chaque flux.

B. Calcul du RCE flux par flux

Pour calculer le résultat de couverture économique global sur notre portefeuille, nous considérons, comme on l'a vu plus haut, que toutes les autres problématiques sont d'ordre comptable, et qu'en conséquence le résultat de couverture global sur une période donnée est la

somme algébrique des résultats de couverture calculés pour chaque flux enregistré en gestion (dit flux passif) sur la dite période.

Pour un flux donné, nous avons besoin d'identifier :

1. la valeur liquidative à laquelle a été passé l'ordre à l'actif (VL actif).
2. la valeur liquidative avec laquelle le nombre de parts à vendre ou acheter a été estimé (VL de saisie), puisqu' au jour de saisie nous ne connaissons pas la VL de conversion (VL d'effet ou VL passif) qui engage la compagnie d'assurances vis-à-vis de son client. Cette valeur liquidative est la dernière connue au moment de la saisie de l'acte (c'est en général celle de J-2 pour une saisie à J.)
3. La valeur d'effet ou VL actif.

Elle nous est indispensable à deux titres :

- Elle nous permet de connaître le nombre de parts réel correspondant à la matérialisation du flux sur l'épargne du client. La différence entre ce dernier et le nombre de parts estimé de l'ordre nous donne le flux d'ajustement de l'actif en UC sur le passif en UC de la compagnie sur cet acte de gestion.
 - Elle nous donne le montant auquel le flux a été réalisé au passif. La VL passif est contractuellement déterminée à réception de toutes les pièces juridiquement nécessaires à la réalisation de l'acte dans les services clients (voir annexes pour un extrait des CG du produit étudié)
4. La valeur liquidative d'ajustement. Comme on l'a vu plus haut, le comité actif passif mensuel va procéder au réadossement de l'actif en UC sur le passif en UC sur les supports dont les écarts dépassent 50 K€. Une étude est actuellement en cours pour supprimer ce plancher de 50 K€. On ne prendra donc aucun plancher en compte. Là aussi, nous considérerons que ce réajustement global réalisé à partir de deux écarts de position à la fin d'un mois donné, est la somme algébrique des flux de réajustement qu'il faudrait réaliser sur chaque ordre.
Ce flux individualisé existe bel et bien, d'ailleurs de façon automatique dans le carnet d'ordres d'autres portefeuilles et/ou autres systèmes AXA voir d'autre sociétés d'assurances (GENERALI etc....). Lorsqu'il est généré par le système, l'ajustement est automatiquement réalisé une fois la VL passif connue. Il est alors très proche de la saisie de l'acte et de la réalisation de l'ordre (deux ou trois jours pour une cotation quotidienne de l'OPCVM). Dans notre cas, le comité se réunissant aux alentours du 25^e jour du mois M+2 concernant l'analyse des positions actif et passif à fin M, nous prendrons cette date et la valeur liquidative correspondante comme variables d'ajustement (ou le jour précédent ou d'après si férié ou non ouvré).

Une fois toutes les valeurs liquidatives connues, deux types d'ordre vont être donnés selon la connaissance ou non par le système de la valeur passif au moment du passage de l'ordre.

- Il connaît la Valeur Liquidative (VL) passif. On est vraisemblablement dans un cas de flux rétroactif, nul n'est besoin d'estimer le nombre de parts à passer à l'actif puisque le nombre de parts passif est connu.

On notera alors que le risque financier est alors intégralement porté par la différence entre les VLs passif de conversion du flux sur le contrat du client et la VL avec laquelle la contrepartie est réalisée à l'actif de l'entreprise.

- Il ne connaît pas la VL passif. Il va devoir estimer à la dernière VL connue (VL de saisie), la contrepartie en parts du montant vendu ou acheté par le client. On distingue alors deux cas :
 - L'ordre en parts est alors passé le même jour que sa date d'effet juridique dans le système et le risque financier est transférée sur l'écart entre le nombre de parts estimé à l'actif et le nombre de parts réel passif du flux (ce dernier est connu à posteriori).

C'est cet écart qui sera corrigé par le flux d'ajustement réalisé à la VL d'ajustement.
 - L'ordre en parts n'est pas passé le même jour que sa date d'effet juridique et le risque financier existe en nombre de parts et en valeur liquidative (estimation du nombre de parts à l'actif et Valeurs liquidatives de réalisation différentes à l'actif et au passif)

Le solde financier de toutes ces opérations financières (achat, vente et correction) nous donne le résultat de couverture sur le flux étudié.

Illustrons par des exemples le calcul de notre résultat de couverture par flux :

soit la réorientation d'épargne en sortie réalisée sur un contrat de notre portefeuille sur le support suivant DWS EMERGING MARKETS pour un montant de 2000 euros brut soit 1990 euros net des frais d'arbitrage. Elle a été saisie le 17/03/2008 en date d'effet du 17/03/2008.

Nous ne connaissons pas la VL d'effet le jour de la saisie, le système va donc chercher la dernière VL connue ce jour la, en l'occurrence celle du 14/03/2008 soit 91.22. L'ordre en parts qui va remonter pour être exécuté le lendemain 18/03/2008 est de 21.81539 parts au titre de cette réorientation d'épargne en sortie.

L'ordre de vente sera réalisé chez le dépositaire à la VL de 85.72 (VL du 18/03/2008)

Du point de vue du client, les 1990 euros seront convertis à la VL du 17/03/2008 soit 86.46.

Soit $1990/86.46=23.01642$ parts.

A posteriori, on constate que l'assureur a vendu 21.81539 parts au prix de 85.72 au lieu de 23.01642 parts au prix de 86.46, soit 1870 euros récupérés au lieu de 1990.

A supposer que le flux de réajustement théorique correspondant (vente de 1.2 parts) a été réalisé le 26/05/2008 (VL de 99.83), il récupèrera 119.89. Soit un résultat de $1807+119.89-1990=0.856$ € et le résultat de couverture est donc de :

$$((-1990/91.22)-(-1990/86.46))*99.83-(((1990/91.22)*85.72)-((1990/86.46)*86.46))=0.856 \text{ €}$$

C. Analyse des résultats 2008

Le résultat de couverture économique global sur l'année serait de **1 418 439 €**.

On constate que la moyenne empirique du résultat de couverture pour un flux est de 15.88 € en 2008, mais l'écart type est à hauteur de 6013 €, donc la dispersion est conséquente. La médiane est nulle, ce qui laisse à penser que les chances d'obtenir un RCE positif sont aussi grandes que celles d'obtenir un RCE négatif.

Les flux financiers à RCE nul s'avèrent être des flux compensés exactement dans les mêmes conditions (dates de saisie d'effet support...) par des montants de sens contraire : le solde des sommes financières engagées est égale à 0 entraînant la nullité du RCE.

Ils sont au nombre de 2241.

L'intervalle interquartile est de 19.44.

50 % des RCE se trouvent dans l'intervalle [-8.8€ ; +10.6€].

99% des RCE se trouvent dans l'intervalle [-2000€ ; +2000€].

Par contre, la valeur de l'étendue (1.7M€) nous laisse à penser qu'un certain nombre de flux financiers peuvent avoir des impacts dépassant le million d'euros. Nous analyserons les tenants et aboutissants de ces cas de figure et notamment les observations dont les RCE maximum et du minimum.

La procédure UNIVARIATE
Variable : rce
Moments

| | | | |
|-----------------|------------|-----------------|------------|
| N | 89297 | Somme poids | 89297 |
| Moyenne | 15.8845159 | Somme obs. | 1418439.62 |
| Écart-type | 6013.07891 | Variance | 36157118 |
| Skewness | 44.6199409 | Kurtosis | 16701.2288 |
| SS non corrigée | 3.22871E12 | SS corrigée | 3.22869E12 |
| Coeff Variation | 37854.9712 | Moy. erreur std | 20.1223393 |

Mesures statistiques de base

| Position | | Variabilité | |
|----------|----------|--------------------------|----------|
| Moyenne | 15.88452 | Écart-type | 6013 |
| Médiane | 0.00000 | Variance | 36157118 |
| Mode | 0.00000 | Étendue | 1784793 |
| | | Intervalle interquartile | 19.44635 |

Tests de position : $\mu_0=0$

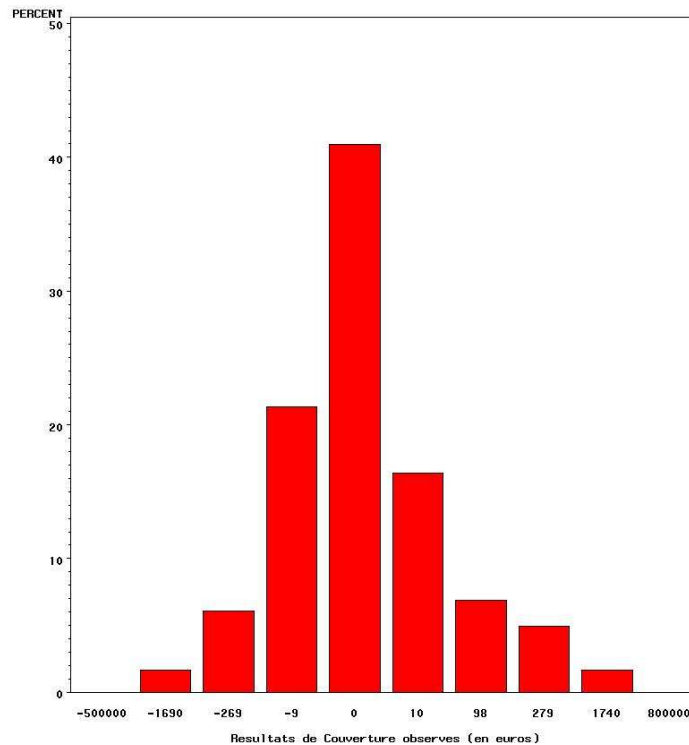
| Test | -Statistique- | -----p Value----- |
|--------------|---------------|-------------------|
| t de Student | t 0.789397 | Pr > t 0.4299 |
| Signe | M 358.5 | Pr >= M 0.0136 |
| Rangs signés | S 31682833 | Pr >= S <.0001 |

Quantiles (Définition 5)

| Quantile | Estimation |
|-------------|--------------|
| 100% Max | 1.03960E+06 |
| 99% | 1.74157E+03 |
| 95% | 2.79176E+02 |
| 90% | 9.82790E+01 |
| 75% Q3 | 1.06494E+01 |
| 50% Médiane | 0.00000E+00 |
| 25% Q1 | -8.79693E+00 |
| 10% | -8.90714E+01 |
| 5% | -2.69237E+02 |
| 1% | -1.69110E+03 |
| 0% Min | -7.45190E+05 |

Ci-dessous la distribution empirique du résultat de couverture 2008 réalisée à partir de la procédure GCHART de SAS.

Frequence en pourcentage des Resultats de Couverture observes



1. Etude des valeurs extrêmes

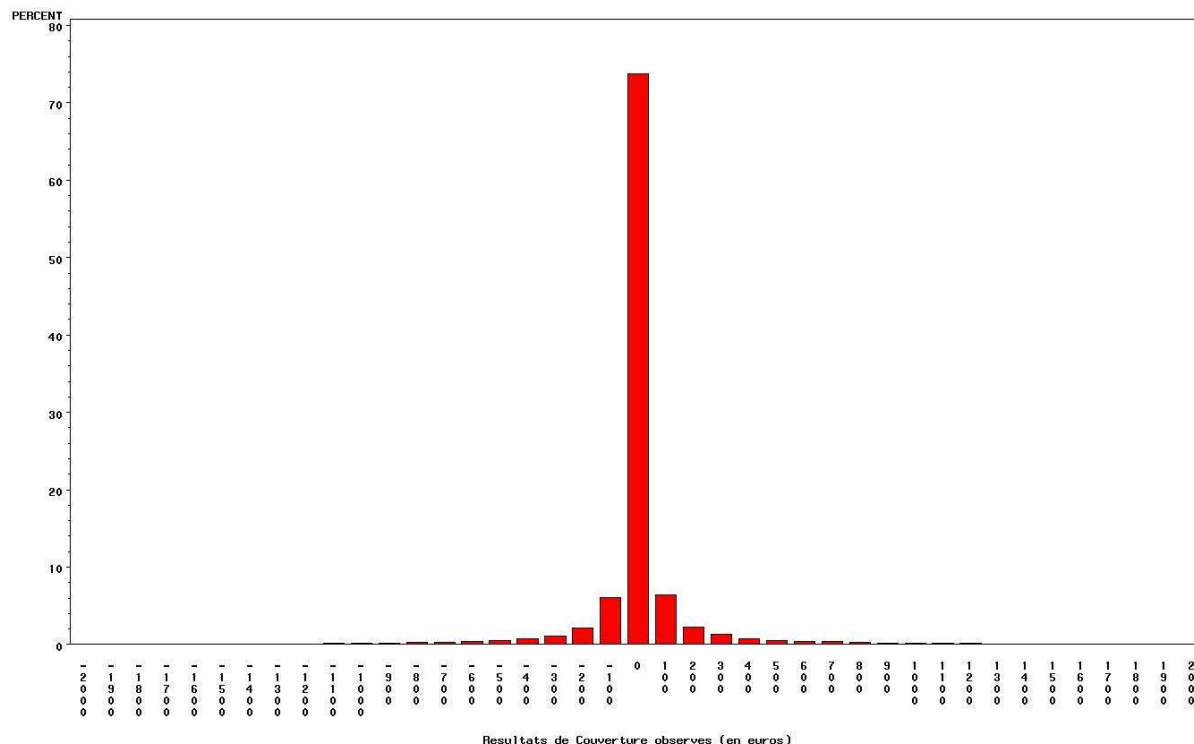
La perte maximale calculée de 745 000 € a été réalisée sur la vente d'un titre à cotation continue effectuée par l'intermédiaire d'un arbitrage sortie de 2 Millions d'euros sur 2 contrats avec un décalage de 188 jours entre la saisie du 04/11/08 et la date d'effet du 30/04/08. Le cours de l'action concernée avait fondu de 40% au 04/11/08 par rapport au 30/04/08. L'explication de cet écart a été retrouvée dans le dysfonctionnement d'un batch informatique (l'arbitrage a été saisi sur internet par le commercial). Le processus d'arbitrage par le net n'a pas fonctionné correctement au niveau des systèmes d'information, le flux n'est remonté que le 04/11/08.

Il faut noter que ce flux a été annulé le 12/11/08 par les équipes de gestion au regard du risque couru et suite à négociation avec le partenaire, flux qui a généré un RCE de 1.039M€ venant heureusement compenser cette perte qui correspond à notre maximum d'ailleurs.

Finalement cela s'est soldé par un gain de 250K€ car entre le 04/11/08 et le 12/11/08, l'action a encore perdu près du quart de sa valeur (on a vendu puis racheté le même nombre de parts). Ce cas de figure nous montre qu'il existe un risque potentiel de dérapage du résultat de couverture en cas d'effet conjugué d'une grande variation des cours de nos supports et d'un retard conséquent entre la date de valeur passif de nos flux et la date à laquelle les flux donnent lieu à des ordres à l'actif.

La même distribution abstraction faite des valeurs extrêmes, dans l'intervalle [-2000€ ; +2000€]

Fréquence en pourcentage des Résultats de Couverture observés dont la valeur absolue est inférieure à 2000 Euros



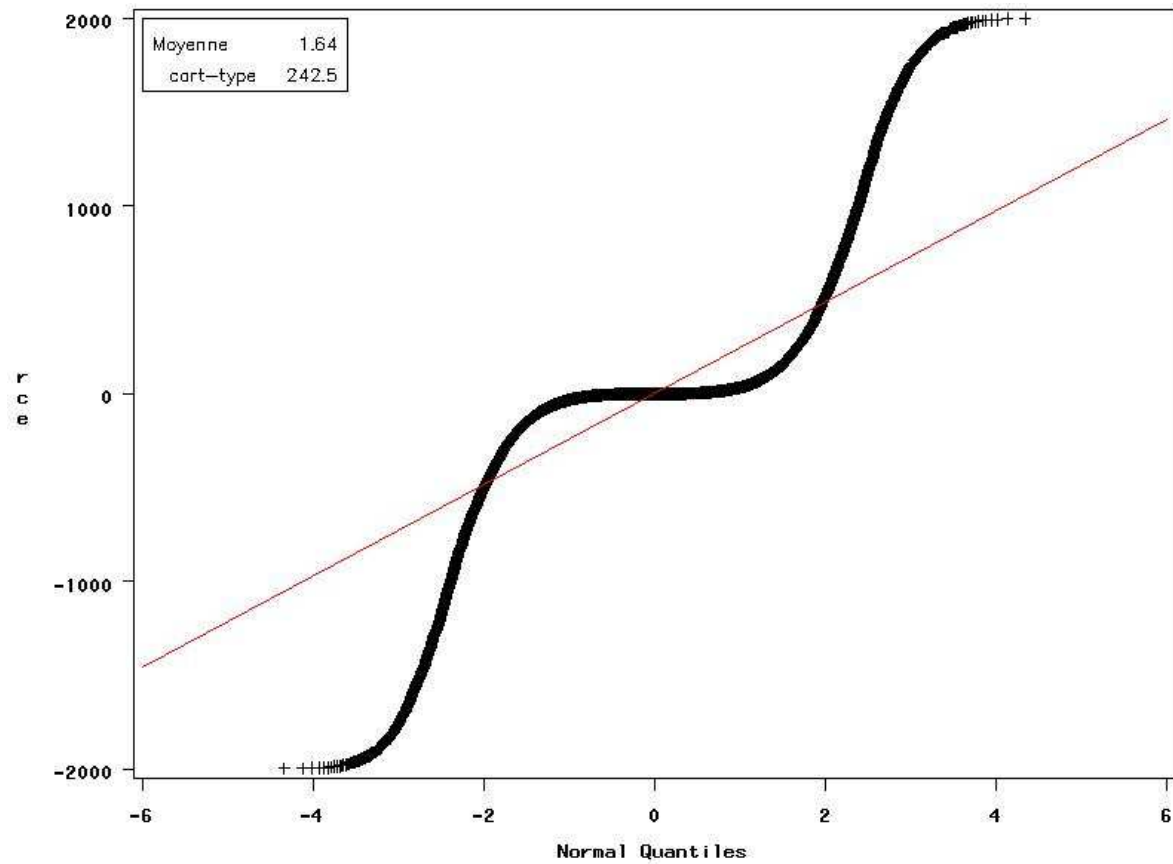
On est tenté de tester la normalité au regard de la forme générale de la distribution.

Les tests de Kolmogorov-Smirnov, Cramer-von Mises, et Anderson-Darling rejettent tous l'hypothèse de normalité (à 5%) comme on peut le constater en lisant les seuils retenus par la proc univariate :

Tests de normalité

| Test | -Statistique-- | -----p Value----- |
|--------------------|----------------|-------------------|
| Kolmogorov-Smirnov | D 0.436713 | Pr > D <0.0100 |
| Cramer-von Mises | W-Sq 6733.457 | Pr > W-Sq <0.0050 |
| Anderson-Darling | A-Sq 31479.96 | Pr > A-Sq <0.0050 |

Par ailleurs la courbe des quantiles empiriques rapprochés des quantiles théoriques de la loi normale montre une courbe sinusoïdale ne correspondant pas à une liaison de type linéaire :



La procédure UNIVARIATE
Variable : rce

Moments

| | | | |
|-----------------|------------|-----------------|------------|
| N | 87789 | Somme poids | 87789 |
| Moyenne | 1.63795126 | Somme obs. | 143794.103 |
| Écart-type | 242.506719 | Variance | 58809.5089 |
| Skewness | 0.02322801 | Kurtosis | 24.0580045 |
| SS non corrigée | 5163004695 | SS corrigée | 5162769168 |
| Coeff Variation | 14805.4906 | Moy. erreur std | 0.81847181 |

Mesures statistiques de base

| Position | | Variabilité | |
|----------|----------|--------------------------|-----------|
| Moyenne | 1.637951 | Écart-type | 242.50672 |
| Médiane | 0.000000 | Variance | 58810 |
| Mode | 0.000000 | Étendue | 3996 |
| | | Intervalle interquartile | 18.23060 |

Tests de position : $\mu_0=0$

| Test | -Statistique- | -----p Value----- | |
|--------------|---------------|-------------------|--------|
| t de Student | t 2.001231 | Pr > t | 0.0454 |
| Signe | M 342.5 | Pr >= M | 0.0174 |

Rangs signés S 30332054 Pr >= |S| <.0001

Quantiles (Définition 5)

| Quantile | Estimation |
|-------------|-------------|
| 100% Max | 1999.51718 |
| 99% | 920.20511 |
| 95% | 234.36591 |
| 90% | 87.09528 |
| 75% Q3 | 9.96705 |
| 50% Médiane | 0.00000 |
| 25% Q1 | -8.26355 |
| 10% | -79.02330 |
| 5% | -221.99110 |
| 1% | -893.19318 |
| 0% Min | -1996.72412 |

Observations extrêmes

| ---Inférieures--- | | ---Supérieures--- | |
|-------------------|-------|-------------------|-------|
| Valeur | Obs. | Valeur | Obs. |
| -1996.72 | 82991 | 1988.23 | 55832 |
| -1996.61 | 72590 | 1990.32 | 59728 |
| -1994.12 | 68782 | 1992.50 | 11879 |
| -1993.70 | 12908 | 1997.25 | 52890 |
| -1993.37 | 91176 | 1999.52 | 29107 |

On notera que confinée à l'intervalle $[-2000€ ; +2000€]$, la moyenne du RCE est de 1.64€ et que ce dernier est nettement moins dispersé : l'écart type est de 242.

Le résultat de couverture économique global sur l'année serait de **143 794 €**.

Nous allons maintenant analyser notre RCE en distinguant les flux d'entrée (souscriptions, versements, réorientation d'épargne) et les flux de sortie (rachats, décès, réorientations d'épargne).

2. Etude du RCE par classe d'actifs

Nous avons segmenté notre RCE par type d'actifs en reprenant la classification standard établie par l'autorité des marchés financiers.

| Famille AMF | Classe Actifs | Nombre de flux | RCE | RCE Moyen | RCE standardisés sur les retards | Volatilité moyenne 2008 pondérée par les montants |
|--------------------|---------------|----------------|-----------|-----------|----------------------------------|---|
| Actions françaises | Actions | 29 | 284 091 € | 9 796 € | - 12 € | 2 |
| Autres | Autres | 543 | 3 516 € | 6 € | - 5 € | 2 |

| | | | | | | |
|--|-------------|---------------|--------------------|-------------|------------|-----------|
| BMTN | Obligations | 5 | 108 889 € | 21 778 € | 22 € | - |
| OPCVM Actions de pays de la zone euro | Actions | 3 980 | 152 967 € | 38 € | 91 € | 30 |
| OPCVM Actions des pays de la communauté européenne | Actions | 14 923 | 317 998 € | 21 € | 3 € | 16 |
| OPCVM Actions européennes | Actions | 3 124 | - 28 408 € | - 9 € | - 14 € | 17 |
| OPCVM Actions françaises | Actions | 8 576 | - 57 907 € | - 7 € | - 78 € | 13 |
| OPCVM Actions internationales | Actions | 26 826 | 435 642 € | 16 € | 277 € | 21 |
| OPCVM Diversifiés | diversifié | 27 022 | - 94 363 € | - 3 € | - 298 € | 8 |
| OPCVM Fonds à Formule | diversifié | 1 291 | - 8 568 € | - 7 € | - 14 € | 14 |
| OPCVM Monétaire euro | Monétaire | 2 404 | - 747 € | - 0 € | - 27 € | 7 |
| OPCVM Monétaire à vocation internationale | Monétaire | 15 | 5 € | 0 € | 0 € | 1 |
| OPCVM Obligation et/ou titres de créances libellés en euro | Obligations | 1 201 | 168 814 € | 141 € | 64 € | 30 |
| OPCVM Obligations et autres titres de créance internationaux | Obligations | 1 232 | 9 561 € | 8 € | 5 € | 5 |
| OPCVM de fonds alternatifs | Obligations | 310 | 146 191 € | 472 € | 90 € | 11 |
| Obligation indexée | Obligations | 57 | - 19 242 € | - 338 € | - 88 € | 34 |
| Total | | 91 538 | 1 418 440 € | 15 € | - € | 13 |

Le RCE standardisé sur les retards consiste à centrer et réduire le montant du RCE sur chaque sous-population définie par une occurrence de retard.

On constate ainsi que le RCE moyen élevé sur la modalité « BMTN » est en fait plus attribuable à un problème de retard. Sur les deux supports concernés, les retards de saisie sont supérieurs à 20 jours.

La volatilité est celle constatée sur les valeurs liquidatives d'effet des flux 2008 (une seule date d'effet pour les BMTN). Il n'a pas été possible de récupérer simplement l'ensemble des valeurs liquidatives de l'année 2008 pour tous les supports.

Si on regroupe par classe d'actifs, la synthèse est la suivante :

| Classe d'actifs | RCE | RCE standardisés | Volatilité moyenne 2008 pondérée par les montants |
|----------------------|--------------------|------------------|---|
| Actions | 1 104 382 € | 260 € | 16 |
| Autres | 3 516 € | 5 € | 2 |
| Monétaire | - 742 € | 28 € | 4 |
| Obligations | 414 214 € | 84 € | 16 |
| diversifié | - 102 931 € | 312 € | 11 |
| Total général | 1 418 440 € | 0,00 € | 13 |

Si on atténue l'effet « retard », les RCE les plus élevés sont observables sur les classes d'actifs les plus volatiles, ce qui constitue un résultat attendu.

3. Distribution RCE sur les entrées

Nous appliquons la procédure univariante **sur les flux d'entrée** pour constater que la moyenne du RCE est alors de 37.5 € avec un écart type de 2518 €.

L'observation des quantiles montre que le RCE est plutôt positif, le quantile à 5% étant de -310 €. Pour mémoire le quantile à 5% du RCE total est de -2700 €.

La procédure UNIVARIATE
Variable : rce

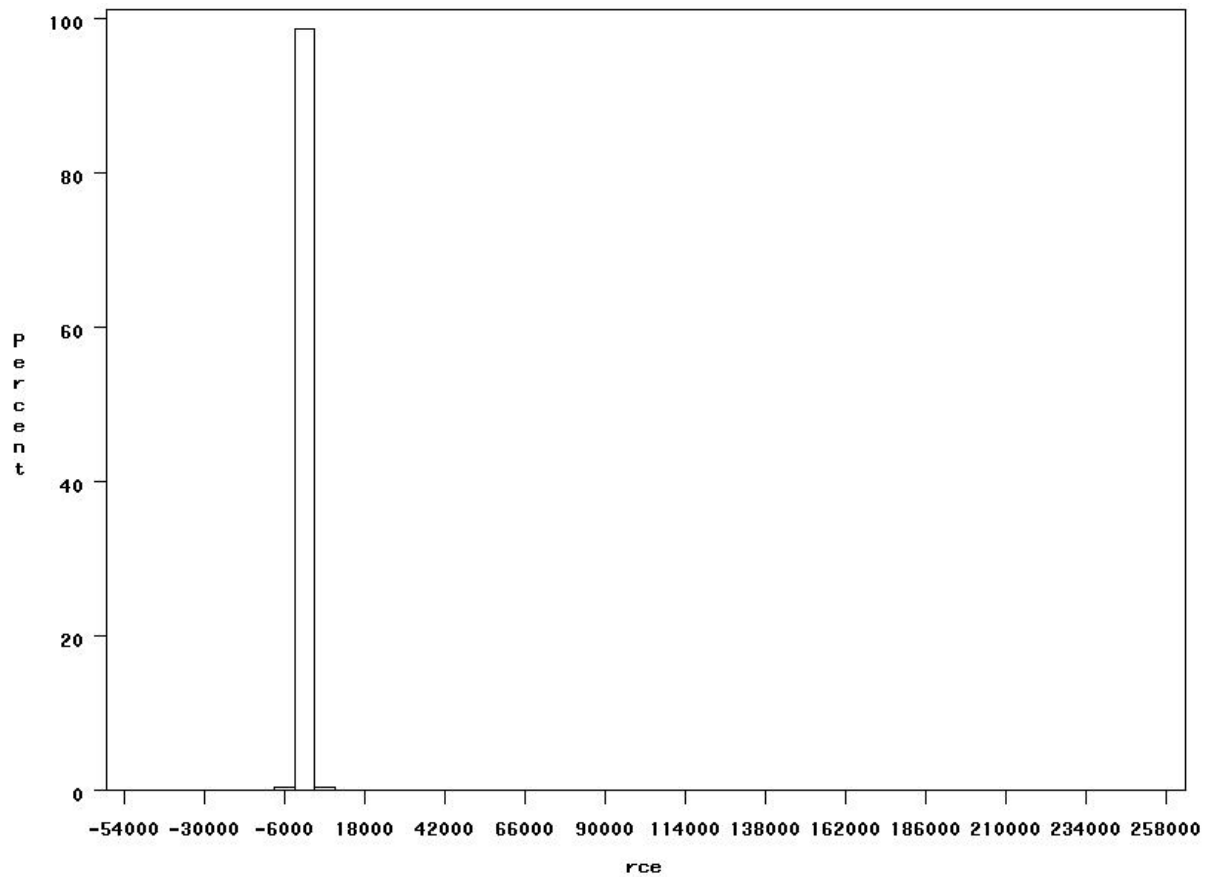
Moments

| | | | |
|-----------------|------------|-----------------|------------|
| N | 38571 | Somme poids | 38571 |
| Moyenne | 37.4961293 | Somme obs. | 1446263.2 |
| Écart-type | 2517.85046 | Variance | 6339570.93 |
| Skewness | 56.0590567 | Kurtosis | 5378.90558 |
| SS non corrigée | 2.44571E11 | SS corrigée | 2.44517E11 |
| Coeff Variation | 6714.96099 | Moy. erreur std | 12.8203382 |

Mesures statistiques de base

| Position | | Variabilité | |
|----------|----------|--------------------------|----------|
| Moyenne | 37.49613 | Écart-type | 2518 |
| Médiane | 0.00000 | Variance | 6339571 |
| Mode | 0.00000 | Étendue | 315512 |
| | | Intervalle interquartile | 28.35616 |

| Quantile | Estimation | 95% Intervalle de confiance sous hypothèse de normalité | |
|-------------|-------------|---|----------|
| 100% Max | 260512.7566 | | |
| 99% | 2299.3890 | 5846.82 | 5943.57 |
| 95% | 351.2875 | 4140.67 | 4217.75 |
| 90% | 123.2561 | 3230.51 | 3298.34 |
| 75% Q3 | 15.0703 | 1708.01 | 1763.69 |
| 50% Médiane | 0.0000 | 12.37 | 62.62 |
| 25% Q1 | -13.2859 | -1688.70 | -1633.02 |
| 10% | -108.8475 | -3223.34 | -3155.52 |
| 5% | -310.4380 | -4142.76 | -4065.67 |



4. Distribution RCE sur les sorties

La procédure UNIVARIATE
Variable : rce

Moments

| | | | |
|-----------------|------------|-----------------|------------|
| N | 50726 | Somme poids | 50726 |
| Moyenne | -0.5485075 | Somme obs. | -27823.59 |
| Écart-type | 7670.0528 | Variance | 58829710 |
| Skewness | 36.3457863 | Kurtosis | 11058.2068 |
| SS non corrigée | 2.98414E12 | SS corrigée | 2.98414E12 |
| Coeff Variation | -1398349.8 | Moy. erreur std | 34.0551694 |

Quantile Estimation sous hypothèse de normalité

| | | | |
|-------------|--------------|---------|---------|
| 100% Max | 1.03960E+06 | | |
| 99% | 1.39948E+03 | 17714.9 | 17971.9 |
| 95% | 2.34180E+02 | 12513.7 | 12718.5 |
| 90% | 8.02978E+01 | 9739.3 | 9919.5 |
| 75% Q3 | 7.78633E+00 | 5099.1 | 5247.0 |
| 50% Médiane | 0.00000E+00 | -67.3 | 66.2 |
| 25% Q1 | -6.03925E+00 | -5248.1 | -5100.2 |

| | | | |
|-----|--------------|----------|----------|
| 10% | -7.47186E+01 | -9920.6 | -9740.4 |
| 5% | -2.42240E+02 | -12719.6 | -12514.8 |

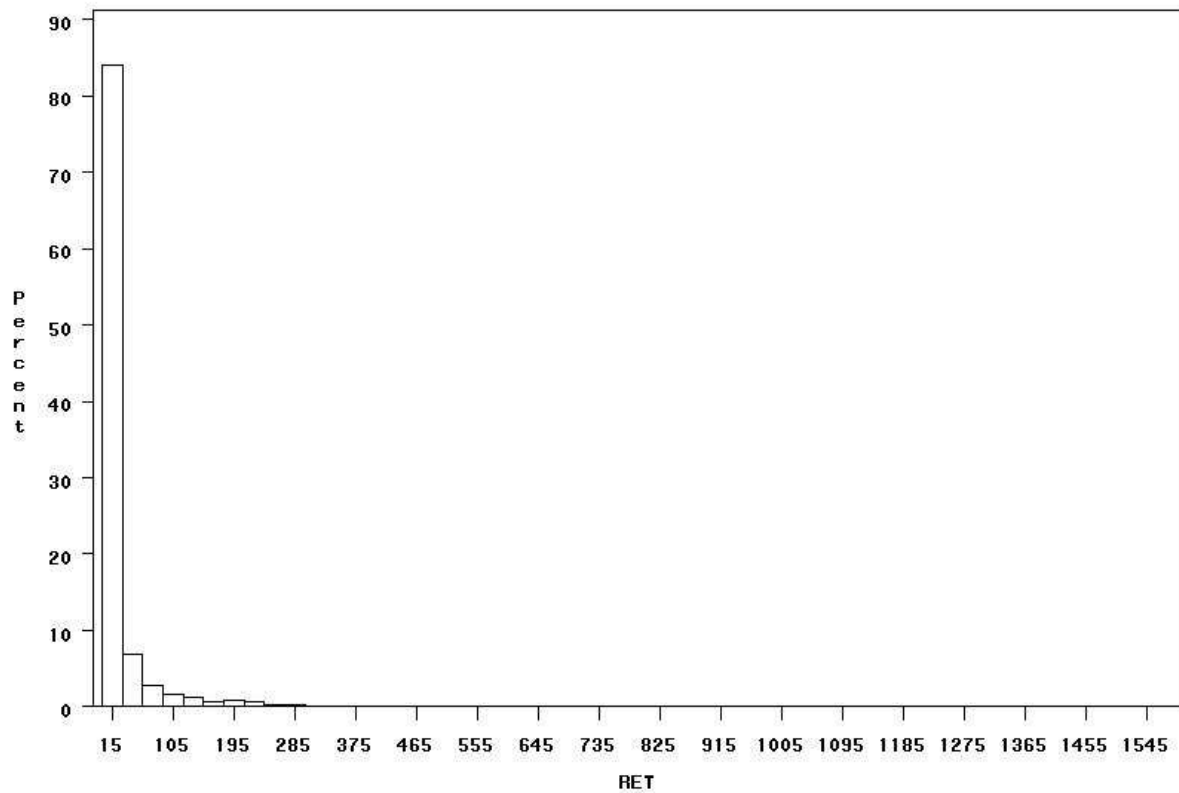
La moyenne est de -0.54 € et l'écart type est de 7670 €. L'observation des quantiles montre que le RCE est plutôt négatif le quantile à 99% étant de -310 €. Pour mémoire le q5% du RCE total est de -2700 €.

On constate donc que les cash-flows positifs ont tendance à générer du RCE positif.

5. Distribution des retards de saisie

Les figures suivantes correspondent à la représentation des retards de saisie en pourcentage (histogramme) et en effectif (courbe).

distribution des retards de saisie hors flux programmes (saisis en avance)



On voit bien que plus de 80% des retards de saisie se trouvent dans la tranche [0 jours, 15 jours]

Puis les effectifs décroissent très rapidement à mesure que les retards s'allongent. Quelques rares grands retards supérieurs à un an sont observés

6. Impact des actes rétroactifs

Nous allons maintenant constater le poids des actes rétroactifs dans les RCE négatifs, nous calculons les RCE moyens et totaux par tranche de retard de saisie.

Nous constatons que la quasi-totalité des actes de gestion générant des résultats de couverture négatifs sont des actes à retard de saisie vis-à-vis de la date d'effet, et l'impact est d'autant plus élevé que le retard est important comme on peut le constater sur la lecture des RCE moyens. On a par ailleurs centré et réduit les RCEs après avoir discrétisé l'espace par classes de valeur absolue de montant. On voit nettement que les RCE sont statistiquement des pertes au-delà 6 mois de retard.

| Délai entre la saisie et l'effet | RCE | Nombre de flux | RCE Moyen | RCE standardisés Par classe de montant |
|----------------------------------|---------------------|----------------|--------------|--|
| saisie en avance | - 725 433 | 10 498 | - 69 | 1 719 |
| pas de retard | - 685 857 | 13 551 | - 51 | 4 337 |
| 1 jours< <15 jours | - 3 295 949 | 11 692 | - 282 | 144 |
| 15 jours< <1 mois | - 974 841 | 1 999 | - 488 | 934 |
| 1 mois< <2 mois | - 904 334 | 1 596 | - 567 | - 1029 |
| 2 mois< <3 mois | - 265 745 | 712 | - 373 | 597 |
| 3 mois< <4 mois | - 412 373 | 426 | - 968 | 581 |
| 4 mois< <5 mois | - 380 066 | 332 | - 1 145 | 510 |
| 5 mois< <6 mois | - 134 999 | 173 | - 780 | - 375 |
| 6 mois< <7 mois | - 1 992 528 | 224 | - 8 895 | - 557 |
| 7 mois< <8 mois | - 63 184 | 159 | - 397 | - 298 |
| 8 mois< <9 mois | - 248 546 | 114 | - 2 180 | - 351 |
| 9 mois< <10 mois | - 161 702 | 120 | - 1 348 | - 281 |
| 10 mois< <11 mois | - 11 805 | 33 | - 358 | - 81 |
| 11 mois< <12 mois | - 42 328 | 28 | - 1 512 | - 118 |
| >12 mois | - 166 494 | 114 | - 1 460 | - 200 |
| Total | - 10 466 186 | 41 771 | - 251 | 0 |

Nous pouvons aussi calculer le rapport relatif du montant brut concerné par chaque flux avec le RCE qu'il a généré.

Nous constatons que le taux moyen est de 1% mais que ce taux a tendance à s'élever très rapidement. Ceci confirme notre remarque précédente.

| Délai entre la saisie et l'effet | RCE | Nombre de flux | (RCE /Montant) Moyen |
|---|---------------------|-----------------------|-----------------------------|
| saisie en avance | - 725 433 | 10 498 | 1 % |
| pas de retard | - 685 857 | 13 551 | 0.8 % |
| 1 jours< <15 jours | - 3 295 949 | 11 692 | 0.8 % |
| 15 jours< <1 mois | - 974 841 | 1 999 | 4 % |
| 1 mois< <2 mois | - 904 334 | 1 596 | 3 % |
| 2 mois< <3 mois | - 265 745 | 712 | 3 % |
| 3 mois< <4 mois | - 412 373 | 426 | 10 % |
| 4 mois< <5 mois | - 380 066 | 332 | 12 % |
| 5 mois< <6 mois | - 134 999 | 173 | 17 % |
| 6 mois< <7 mois | - 1 992 528 | 224 | 16 % |
| 7 mois< <8 mois | - 63 184 | 159 | 17 % |
| 8 mois< <9 mois | - 248 546 | 114 | 18 % |
| 9 mois< <10 mois | - 161 702 | 120 | 18 % |
| 10 mois< <11 mois | - 11 805 | 33 | 19 % |
| 11 mois< <12 mois | - 42 328 | 28 | 24 % |
| >12 mois | - 166 494 | 114 | 11 % |
| Total | - 10 466 186 | 41 771 | 1 % |

3eme Partie

Modélisation de la distribution des résultats de
couverture

III. MODELISATION PAR LE PROCESSUS RSLN2

A. Présentation du processus

Le Processus RSLN2 (Regime-Switching Model of Long-Term Stock Returns) a été introduit par Bollen en 1998 pour évaluer le prix des options américaines puis repris par Hardy en 2003 pour modéliser les rendements mensuels de TSE 300 et S & P 500. Il peut être considéré comme une alternative au modèle de Black & Sholes.

L'approche originale de Black & Sholes suppose que les rendements des prix des actifs financiers suivent un mouvement brownien géométrique. En notant S_t le prix d'un actif financier à l'instant t , on a sous Black & Sholes :

$$\log\left(\frac{S_{t+1}}{S_t}\right) \sim N(\mu, \sigma^2)$$

Ce modèle est simple et flexible. En outre, il donne des approximations raisonnables quand on l'applique à court terme. Cependant, des études empiriques ont montré que le modèle de Black & Sholes atteint ses limites lorsqu'on approche le mouvement extrême des prix des actifs (à long terme) ainsi que la variation de leurs volatilités.

Une façon simple d'améliorer le modèle de Black & Sholes consiste à se donner k régimes $\rho_1, \rho_2, \dots, \rho_k$, chacun étant décrit par une loi normale $N(\mu_i, \sigma_i^2)$ où le passage d'un régime à un autre est aléatoire. Pour simplifier, on suppose que le processus nous informant du régime en cours à un instant donné est markovien. Autrement dit, les probabilités de changement de régime dépendent uniquement du régime courant et non de l'historique du processus des régimes. Sous cette hypothèse, les changements de régimes sont complètement décrits par une matrice carrée $(P_{i,j})_{1 \leq i, j \leq k}$ de probabilités de transition ($P_{i,j}$ désigne la probabilité de passage du

régime ρ_i au régime ρ_j). Le modèle ainsi obtenu se note RSLN(k). Le modèle de Black & Sholes est un ILN (independent log normal) et le modèle dont il est question dans ce chapitre est un RSLN(2) (noté simplement RSLN2).

Hardy démontre que les modèles Regime-Switching en général et le RSLN2 en particulier appréhendent statistiquement mieux que les autres (Arima, Arch Garch...) les périodes de trouble ou incertitude financière (elle utilise pour cela un certain nombre d'indicateurs):

Le marché ou l'actif étudié va passer d'un régime de croisière à volatilité faible (stable) et moyenne haute à un autre de « crise » dont la volatilité est forte (instable) et la moyenne faible (chute des cours). La haute volatilité peut provenir alors de troubles économiques et politiques à court terme.

En ce qui nous concerne, nous allons utiliser le processus RSLN2 pour modéliser la série chronologique des rendements journalier du CAC 40 depuis le 1 mars 1990 jusqu'au 7 août 2009.

Nous avons choisi le CAC 40 comme actif générique dont l'évolution représente l'ensemble des supports de notre portefeuille étant donné la majorité de fonds cotés sur la place de Paris dans notre portefeuille.

Ce que nous avons comme données sont les cours (S_t) du CAC durant la période étudiée.

La série (Y_t) des logarithmes des rendements journaliers du CAC est donnée par :

$$Y_t = \log\left(\frac{S_{t+1}}{S_t}\right)$$

(logarithme du rendement journalier au jour $t+1$).

L'hypothèse du processus RSLN2 est que $Y_t(0 \leq t \leq n)$ switche entre deux lois normales

$N(\mu_1, \sigma_1^2)$ et $N(\mu_2, \sigma_2^2)$. Soit $P = (p_{ij})_{1 \leq i, j \leq 2}$ la matrice des probabilités de transition, on a :

$$p_{ij} = \Pr[\rho_{t+1} = j | \rho_t = i] \quad (1 \leq i, j \leq 2)$$

(qui est la probabilité de passage du régime i au régime j).

Comme on a visiblement $p_{11} + p_{12} = 1$ et $p_{21} + p_{22} = 1$, il reste 6 paramètres à estimer pour notre modèle ; l'ensemble de ces paramètres est :

$$\Theta = \{\mu_1, \mu_2, \sigma_1, \sigma_2, p_{11}, p_{22}\}.$$

Pour estimer ces paramètres, nous utilisons la méthode du *maximum de vraisemblance*.

B. Calcul du maximum de vraisemblance

La vraisemblance du vecteur aléatoire $\mathbf{y} = (y_1, y_2, \dots, y_n)$ représentant l'ensemble de nos observations est :

$$L(\Theta) = f(y_1 | \Theta) f(y_2 | \Theta, y_1) f(y_3 | \Theta, y_1, y_2) \times \dots \times f(y_n | \Theta, y_1, \dots, y_{n-1}),$$

Où $f(y_i)$ est la fonction de densité marginale de y_i . Notons que f est entièrement déterminée une fois connu le régime suivi à un certain instant, c'est simplement la fonction de densité de l'une des deux lois normales $N(\mu_1, \sigma_1^2)$ ou $N(\mu_2, \sigma_2^2)$.

Pour calculer $L(\Theta)$, nous devons calculer pour tout t :

$$f(y_t | y_{t-1}, \dots, y_1, \Theta).$$

Ce calcul utilise des relations de récurrence qu'on trouvera dans Hamilton & Susmel (1994), reprises aussi dans Hardy (2003). On rappelle ces relations dans ce qui suit :

Formules initiales

$$f(\rho_1 = 1, y_1 | \Theta) = \pi_1 \Phi\left(\frac{y_1 - \mu_1}{\sigma_1}\right)$$

$$f(\rho_1 = 2, y_1 | \Theta) = \pi_2 \Phi\left(\frac{y_1 - \mu_2}{\sigma_2}\right),$$

Où Φ désigne la fonction de densité de la loi normale centrée réduite et π_i désigne la probabilité pour que le processus RSLN2 en question soit au régime i . On montre aisément

que l'on a : $\pi_1 = \frac{p_{21}}{p_{12} + p_{21}}$ et $\pi_2 = \frac{p_{12}}{p_{12} + p_{21}}$.

Ensuite :

$$f(y_1 | \Theta) = f(\rho_1 = 1, y_1 | \Theta) + f(\rho_1 = 2, y_1 | \Theta)$$

$$\Pr(\rho_1 | y_1, \Theta) = \frac{f(\rho_1, y_1 | \Theta)}{f(y_1 | \Theta)}.$$

Formules récurrentes

Pour tout t , en supposant connu $f(y_{t-1}|y_{t-2}, y_{t-3}, \dots, y_1, \Theta)$, alors $f(y_t|y_{t-1}, y_{t-2}, \dots, y_1, \Theta)$ se calcule grâce aux trois relations récurrentes suivantes :

$$\Pr(\rho_{t-1}|y_{t-1}, y_{t-2}, \dots, y_1, \Theta) =$$

$$\frac{f(\rho_{t-1}, \rho_{t-2} = 1, y_{t-1}|y_{t-2}, y_{t-3}, \dots, y_1, \Theta) + f(\rho_{t-1}, \rho_{t-2} = 2, y_{t-1}|y_{t-2}, y_{t-3}, \dots, y_1, \Theta)}{f(y_{t-1}|y_{t-2}, y_{t-3}, \dots, y_1, \Theta)},$$

$$f(\rho_t, \rho_{t-1}, y_t|y_{t-1}, \dots, y_1, \Theta) = \Pr(\rho_{t-1}|y_{t-1}, \dots, y_1, \Theta) \times \Pr(\rho_t|\rho_{t-1}, \Theta) \times f(y_t|\rho_t, \Theta)$$

(remarquez que $\Pr(\rho_t|\rho_{t-1}, \Theta)$ n'est rien d'autre que la probabilité de transition du régime ρ_{t-1} au régime ρ_t et que $f(y_t|\rho_t, \Theta) = \Phi'((y_t - \mu_{\rho_t})/\sigma_{\rho_t})$,

$$f(y_t|y_{t-1}, \dots, y_1, \Theta) = \sum_{\substack{\rho_t \in \{1,2\} \\ \rho_{t-1} \in \{1,2\}}} f(\rho_t, \rho_{t-1}, y_t|y_{t-1}, \dots, y_1, \Theta).$$

Une fonction programmée en VBA qui calcule la valeur du logarithme de vraisemblance pour un ensemble donné Θ de paramètres, et relative à un certain nombre d'observations est donnée en annexe.

Estimation des paramètres

Par un programme VBA, on a défini une fonction $LogL$ à 6 paramètres $p_{11}, p_{22}, \mu_1, \mu_2, \sigma_1, \sigma_2$ qui calcule le logarithme de vraisemblance du modèle RSLN2 correspondant à la série des cours du CAC40. Sur une feuille Excel, on a initialisé ces paramètres à :

$p_{11} = 0.5, p_{22} = 0.5, \mu_1 = \mu_2 = E(Y) \approx 0.013\%, \sigma_1 = \sigma_2 = \sigma(Y) \approx 1.42\%$ (moyenne et écart type empirique constatés sur notre série des log-rendements du CAC 40).

Dans une autre cellule, on a calculé $LogL(p_{11}, p_{22}, \mu_1, \mu_2, \sigma_1, \sigma_2)$. Ensuite, par le solveur Excel, on a maximisé $LogL$ par modification des valeurs initiales de chacun des paramètres en question, tout en imposant les contraintes naturelles :

$$0 \leq p_{11} \leq 1, 0 \leq p_{22} \leq 1, \sigma_1 > 0, \sigma_2 > 0.$$

Le solveur nous a renvoyé alors les solutions suivantes :

$$p_{11} \approx 0.98, p_{22} \approx 0.99, \mu_1 \approx -0.13\%, \mu_2 \approx 0.05\%, \sigma_1 \approx 2.35\%, \sigma_2 \approx 1.01\%.$$

Elles maximisent la vraisemblance de notre série de données.

Le processus RSLN2 ainsi déterminé, la prochaine étape consiste à valider le processus.

C. Comparaison avec d'autres modèles

Les modèles les plus complexes doivent apporter des améliorations significatives en rapport avec leur complexité (c'est le principe de parcimonie). Pour les modèles à complexité identique (i.e., le même nombre de paramètres), il est naturel de choisir celui qui possède la plus forte vraisemblance. Cependant, ce même critère de choix ne peut s'appliquer pour des modèles à complexités différentes car, tout simplement, il ne tient pas compte de ces dernières. Pour remédier à cette problématique, plusieurs critères de sélection ont été mis en place tenant compte de la complexité des modèles.

1. Le test du rapport de vraisemblance (LRT)

Le test de rapport de vraisemblance permet de comparer entre deux modèles, l'un emboîté dans l'autre. Soit M_1 un modèle à k_1 paramètres, emboîté dans un autre modèle M_2 à k_2 paramètres (c'est-à-dire que M_1 est un cas particulier de M_2). On a donc $k_1 < k_2$. Soient ℓ_1 et ℓ_2 les logarithmes des vraisemblances respectives des deux modèles M_1 et M_2 . Le test statistique du critère LRT est :

$$2(\ell_2 - \ell_1).$$

L'hypothèse nulle est :

$$H_0 : \text{Pas d'amélioration significative apportée par } M_2.$$

Sous l'hypothèse nulle, la statistique suit une loi χ^2 de degré de liberté égale à la différence entre le nombre des paramètres des deux modèles (i.e., $k_2 - k_1$). La petite p-value du test entraîne alors le rejet de l'hypothèse nulle.

N.B : Quand les modèles à comparer ne sont pas emboîtés, on peut toujours utiliser les tests LRT mais dans ce cas, la distribution χ^2 en question n'est qu'une approximation.

2. Le critère d'information d'Akaike (AIC)

Le critère d'information d'Akaike retient le modèle qui maximise $\ell - k$, où ℓ désigne le logarithme de vraisemblance du modèle et k désigne le nombre de paramètres du même modèle. Selon ce critère, chaque nouveau paramètre doit contribuer à améliorer (augmenter) d'une unité le logarithme de vraisemblance. Bien que son fondement soit purement heuristique, ce critère est très populaire pour sa simplicité.

3. Le critère de Shwartz-Bayes (SBC)

Le critère de Shwartz-Bayes retient le modèle qui maximise $\ell - 1/2 k \log n$, où n désigne la taille de l'échantillon (i.e., de la série observée). Comme pour le critère d'Akaike, ce critère exige à ce que chaque paramètre supplémentaire contribue à améliorer le logarithme de vraisemblance. La différence avec AIC est que pour SBC, l'amélioration du logarithme de vraisemblance doit croître avec la taille de l'échantillon.

Dans le tableau qui suit, nous allons choisir à la lumière de ces critères, le meilleur modèle représentatif des rendements du CAC.

4. Comparaison entre quelques modèles en utilisant les critères de sélection précédents

| Modèle | Nombre de paramètres (k) | Logarithme de vraisemblance (ℓ) | Critères | |
|--------|------------------------------|--|--------------------|-------------------------------|
| | | | AIC ($\ell - k$) | SBC ($\ell - 1/2 k \log n$) |
| ILN | 2 | 13920.23 | 13918.23 | 13911.73 |
| RSLN2 | 6 | 14545.25 | 14539.25 | 14519.75 |

Par ailleurs, le test du rapport de vraisemblance entre les deux modèles emboîtés ILN (Black & Scholes) et RSLN2 a donné comme p-value 2.3×10^{-269} , ce qui est une valeur très petite. Par conséquent, l'hypothèse nulle est rejetée et le modèle RSLN2 apporte donc une amélioration significative par rapport au modèle de Black & Scholes.

5. Fonction de répartition globale du modèle RSLN2

On peut voir l'ensemble des valeurs de notre échantillon y_1, y_2, \dots, y_n comme des valeurs d'une seule variable aléatoire et non d'un processus. De ce point de vue, on peut calculer la fonction de répartition globale F d'un modèle RSLN2. On a pour tout $x \in R$:

$$\begin{aligned}
 F(x) &= \Pr(Y \leq x) \\
 &= \Pr(Y \leq x | \rho = 1) \cdot \Pr(\rho = 1) + \Pr(Y \leq x | \rho = 2) \cdot \Pr(\rho = 2) \\
 &= \pi_1 \Phi\left(\frac{x - \mu_1}{\sigma_1}\right) + \pi_2 \Phi\left(\frac{x - \mu_2}{\sigma_2}\right),
 \end{aligned}$$

où Φ désigne la fonction de répartition de la loi normale centrée réduite.

On voit donc que F est complètement déterminée en fonction des paramètres du modèle. Ce fait nous permettra de faire d'autres types de tests pour décider s'il convient ou non de retenir le modèle RSLN2 sélectionné. Ces tests s'appuient sur les différences entre les effectifs réels et les effectifs théoriques (résultant de la fonction de répartition du modèle) de notre échantillon d'observation.

6. Le test Khi-Deux de Pearson :

On répartit les valeurs de l'échantillon (de taille n) dans k classes distinctes et on calcule les effectifs o_i de ces classes. D'un autre côté, on calcule les probabilités théoriques p_i d'appartenir à chaque classe (en se servant de la fonction de répartition du modèle) et puis les effectifs théoriques e_i qui en résultent ($e_i = n \cdot p_i$).

Ensuite, on calcule la statistique :

$$Q = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i}$$

qui donne une mesure de l'écart existant entre les effectifs théoriques attendus et ceux observés dans l'échantillon. Par suite, on se donne un seuil de tolérance α (en général $\alpha = 5\%$) et on compare Q à la p-value de la loi χ^2 à $k-1$ degrés de liberté qui est par définition $\chi_{k-1, \alpha}^2 := \Pr(X \geq \alpha)$, avec X suit une loi χ^2 à $k-1$ degrés de liberté.

Si l'on trouve $Q > \chi_{k-1, \alpha}^2$ (avec n suffisamment grand) alors l'hypothèse d'avoir effectivement affaire à la répartition théorique choisie est à rejeter avec une probabilité d'erreur d'au plus α . La même hypothèse peut être retenue dans le cas contraire.

N.B. : Dans le choix des classes, les effectifs théoriques doivent être significatifs ; il est suggéré de vérifier que $np_i(1-p_i) \geq 5$ pour tout i . Le nombre de classes doit aussi être en rapport avec la taille de l'échantillon, il semble que $k \approx \sqrt{n}$ convient bien.

Application

Dans notre cas, on a partagé notre échantillon de taille $n = 4907$ en $k = 92$ classes. On a calculé les effectifs puis les effectifs théoriques (relatifs à la distribution du modèle RSLN2 sélectionné) pour chaque classe. Enfin, on a calculé la statistique Q du test χ^2 et on a trouvé que $Q \approx 88.02$. En regardant par ailleurs la table statistique des p-values de la loi χ^2 , on voit que la p-value correspondante au seuil de tolérance de 5% et au degré de liberté 85 (92-1-6 paramètres estimés) est 107.52. Comme, on a bien $Q < 107.52$, le modèle RSLN2 sélectionné est **retenu** par le test χ^2 .

7. Le test de Kolmogorov-Smirnov

Le test de Kolmogorov-Smirnov est utilisé pour déterminer si un échantillon donné suit ou non une loi donnée par sa fonction de répartition continue. On l'utilise également pour voir si deux échantillons donnés suivent ou non une même loi.

L'idée de ce test consiste à comparer entre les deux fonctions de répartition, l'empirique et la théorique. Une certaine distance entre ces deux fonctions permet de déterminer si la loi choisie convient ou non à l'échantillon d'étude.

Comme pour le test Khi-Deux, on répartit les valeurs de l'échantillon d'étude (de taille n) en k classes disjointes $[x_0, x_1[, [x_1, x_2[, \dots, [x_{k-1}, x_k[$ et on calcule les effectifs o_i de ces classes, puis les probabilités empiriques $\bar{p}_i = \frac{o_i}{n}$ de chacune des classes. Enfin, on calcule les valeurs de la fonction de répartition empirique F_n sur l'ensemble de toutes les k classes (en cumulant simplement les probabilités empiriques \bar{p}_i). On a pour tout réel x :

$$F_n(x) = \begin{cases} \sum_{i=1}^{\ell} \bar{p}_i & \text{si } x \text{ appartient à la } \ell \text{ ème classe} \\ 0 & \text{si } x \text{ est au dessous de toutes les classes} \\ 1 & \text{si } x \text{ est au dessus de toutes les classes.} \end{cases}$$

Par ailleurs, on calcule les valeurs de la fonction de répartition théorique aux points d'extrémité de chaque classe.

On calcule finalement la statistique de Kolmogorov-Smirnov, donnée par :

$$D_n = \sqrt{n} \sup_n |F_n(x) - F(x)|$$

En remarquant que la fonction F_n est une fonction en escalier, continue à droite (mais pas à gauche), on a simplement :

$$D_n = \sqrt{n} \text{Max}_{0 \leq i \leq k+1} (|F_n(x_i) - F(x_i)|, |F_n(x_{i-1}) - F(x_i)|)$$

(En conventionnant $x_{-1} = 0, x_{k+1} = +\infty$).

Un théorème de Kolmogorov-Smirnov affirme que :

$$\Pr(D_n > \alpha) \xrightarrow{n \rightarrow +\infty} \text{KL}(\alpha),$$

Où KL est la fonction de Kolmogorov-Smirnov, qui est définie par :

$$\text{KS}(t) := 2 \sum_{r=1}^{+\infty} (-1)^{r-1} e^{-2r^2 t^2}.$$

Le test de Kolmogorov-Smirnov est basé sur ce théorème. On sait que $\text{KS}(1,36) \approx 0,05$, donc si l'on trouve que $D_n > 1,36$, on pourra rejeter l'hypothèse que notre échantillon suit la loi de fonction de répartition F au risque d'erreur de 5%. Dans le cas contraire, la même hypothèse peut être acceptée.

Une démarche quelque peu différente consiste simplement à calculer la p-value $\text{KS}(D_n)$. L'hypothèse que notre échantillon suit la loi de fonction de répartition F n'est rejetable qu'au risque de $\text{KS}(D_n)$. En fixant un seuil α , la même hypothèse peut donc être retenue dès que $\text{KS}(D_n) < \alpha$. Par conséquent, plus $\text{KS}(D_n)$ est petit, plus la distribution de l'échantillon de données approche la distribution d'étude, idéalement $D_n = 0$.

Il est à noter aussi que la série définissant la fonction KS converge tellement vite qu'une sommation jusqu'à $r = 3$ est suffisante pour commettre une erreur $< 10^{-4}$ dans les calculs.

Application

En utilisant les mêmes classes que pour le test χ^2 , on a trouvé que la statistique de Kolmogorov-Smirnov vaut $D_n \approx 0,26$, ce qui nous donne $\text{KS}(D_n) \approx 2,67 \times 10^{-7}$, qui est visiblement une p-value extrêmement petite. Le modèle RSLN2 sélectionné a donc passé le test de Kolmogorov-Smirnov avec succès. D'ailleurs, on constate dans les calculs que les fonctions de répartition empirique et théorique sont très proches.

IV. ETUDE DE LA STATISTIQUE DES RETARDS

Nous devons calculer les écarts de dates impactant notre calcul de RCE (c'est-à-dire les différences **date d'effet** – **date de saisie** : toutes les autres dates nécessaires à la détermination des VL du RCE étant liées à l'une ou à l'autre) pour toute l'année 2008.

Lorsque cette différence vaut 1, c'est un acte saisi sans retard relatif à sa date d'effet. Plus elle s'éloigne de 1, plus le retard est conséquent. Nous avons déduit de ces délais les jours non travaillés par les équipes du service client qui correspondent aux weekends et aux jours fériés à l'aide d'une macro visual basic.

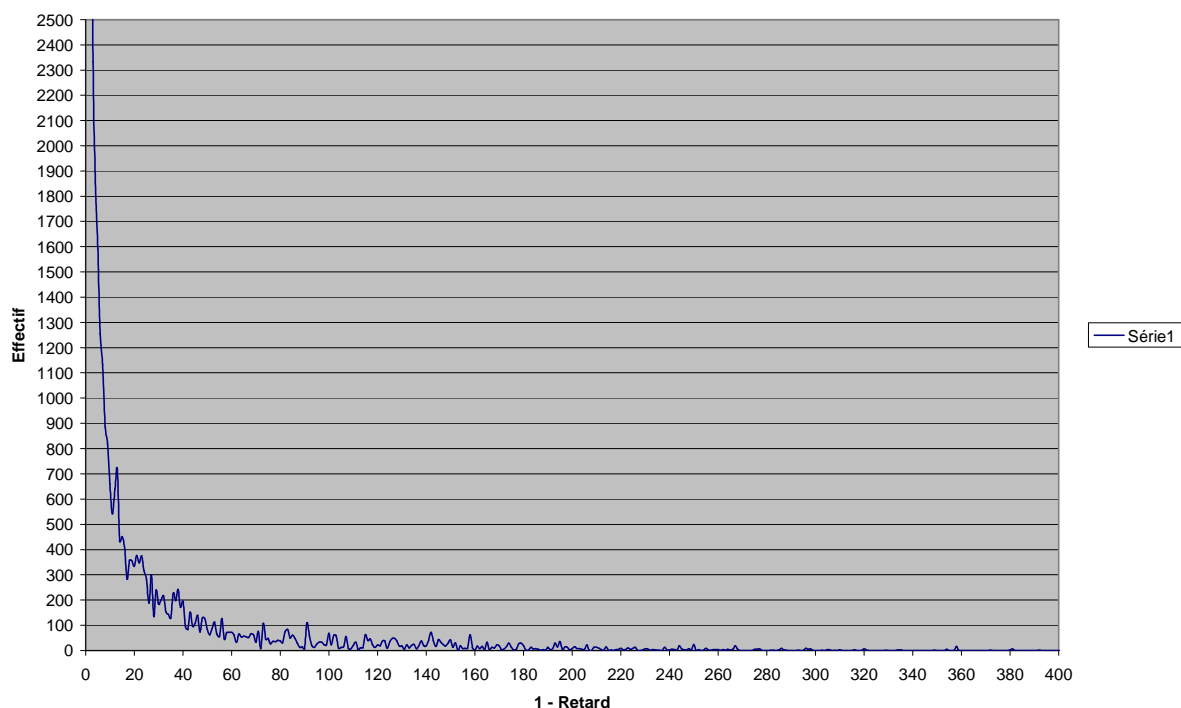
Nous avons travaillé sur 68 625 couples (date d'effet, date de saisie).

Les effectifs correspondants aux différents retards sont naturellement décroissants.

Avant de commencer la modélisation, nous avons ajusté notre variable délai pour la rendre compatible avec l'expression intuitive d'un retard en appliquant la transformation affine :

$$R = 1 - (\text{dated'effet} - \text{datedesaisie}).$$

Ceci nous ramène à une variable aléatoire à valeurs positives et $R=0$ sans retard de saisie, $R=1$ si un jour de retard est observé etc...



A. Loi exponentielle

Le premier modèle qui peut être utilisé pour la nouvelle série statistique R est le modèle de la loi exponentielle. Nous utiliserons la notion de « taux de panne » bien connue en fiabilité (voir bibliographie) dont on démontre qu'il est constant pour une variable aléatoire donnée de loi exponentielle¹ grâce à l'absence de mémoire de la loi.

En ce qui nous concerne, le taux de panne d'un retard de n jours est la probabilité pour qu'un retard n'atteindra pas $(n+1)$ jours sachant qu'il a déjà atteint n jours, il est donc égal à

$$\tau(n) = \Pr(R < n+1 | R \geq n) = \frac{\Pr(R = n)}{\Pr(R \geq n)}.$$

En modélisation du temps de vie d'un organe par exemple, le taux de panne relatif à un âge donné n est la probabilité pour que l'organe en question tombe « en panne » à l'instant qui suit sachant qu'il est en bonne santé à l'instant n .

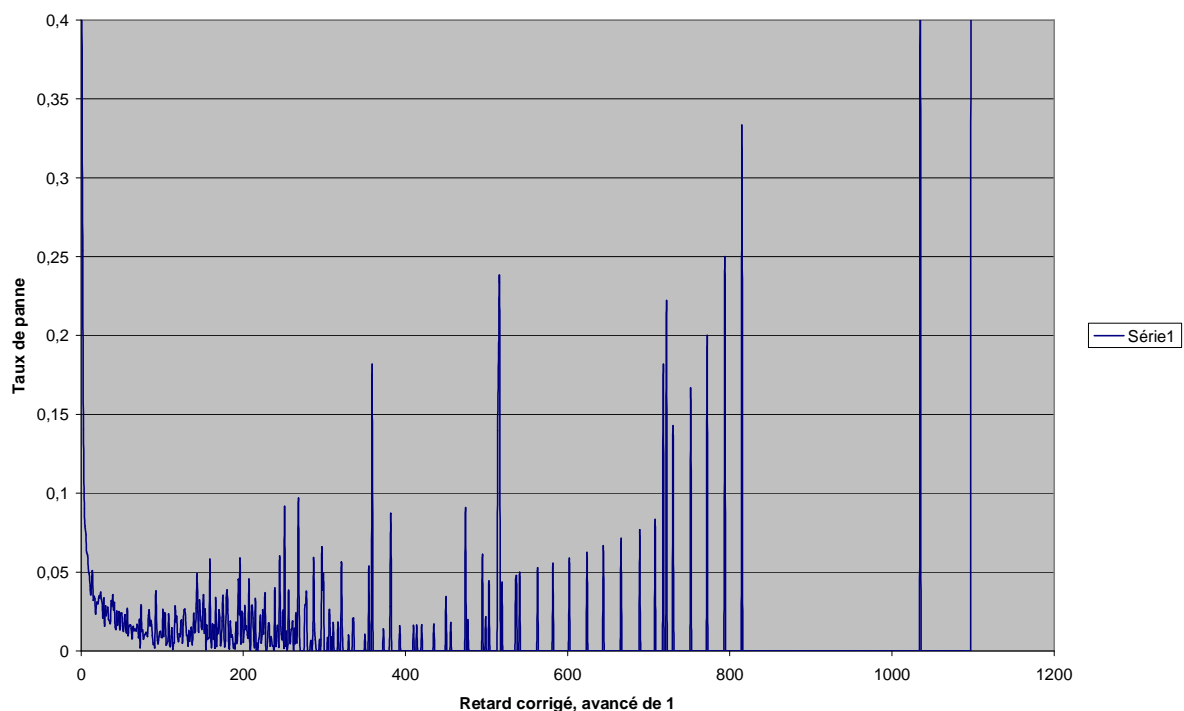
Quand le taux de panne est croissant, on parle de l'effet de vieillissement et quand il est décroissant, on parle du défaut de jeunesse.

¹ Le taux de panne de la loi exponentielle $Exp(\lambda)$ vaut λ .

Pour une loi continue de fonction de densité f , le taux de panne est donné par :

$$\tau(t) = \frac{f(t)}{\int_t^{+\infty} f(x)dx}$$

Nous constatons sur le graphique suivant que les taux de panne observés sur nos retards sont dans un premier temps significativement décroissants puis deviennent plus ou moins stables avec d'importantes perturbations. La loi exponentielle ne convient donc pas pour représenter notre série des retards.



En modélisation statistique, quand un phénomène représentant un temps d'attente (ou de vie) d'un évènement rare possède un taux de panne non stable (i.e., croissant ou décroissant), on cherche à le représenter généralement par une loi de Weibull, une loi Gamma ou une loi mélange d'exponentielles.

Nous allons donc tester successivement ces trois modèles.

B. Loi de Weibull

On appelle loi de Weibull de paramètres k et λ (k et λ étant des réels strictement positifs), la loi dont la fonction de répartition F est de support $[0, +\infty[$ et est donnée par :

$$F(t) = 1 - e^{-(t/\lambda)^k} \quad (\forall x \in [0, +\infty[).$$

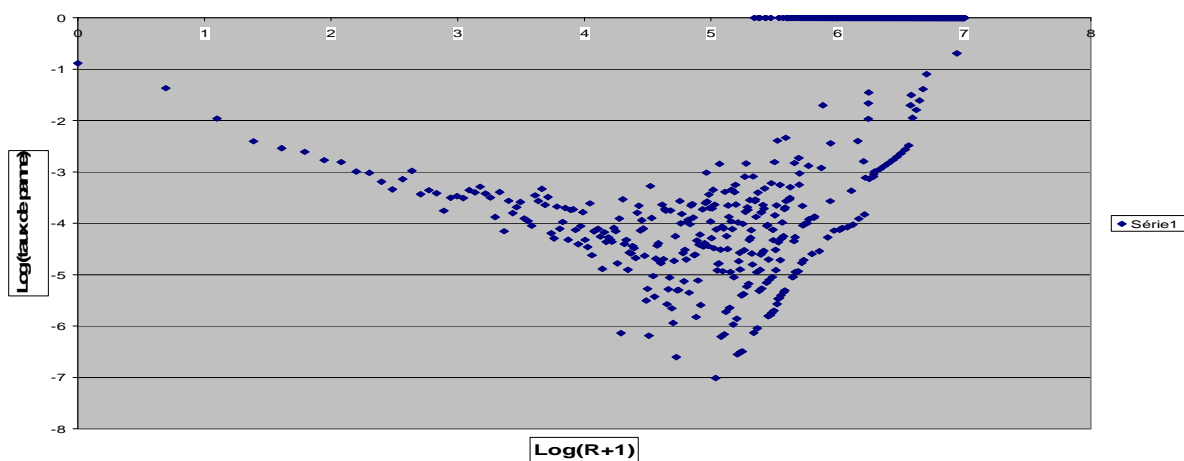
Un calcul simple montre que le taux de panne d'une loi de Weibull de paramètres k, λ est :

$$\tau(t) = \frac{k}{\lambda} \left(\frac{t}{\lambda} \right)^{k-1}.$$

La loi de Weibull (de paramètres k, λ) a donc un taux de panne croissant si $k > 1$ et décroissant si $k < 1$. On retrouve la loi exponentielle de paramètre $1/\lambda$ dans le cas $k = 1$.

La loi de Weibull a un double avantage par rapport à la loi exponentielle. Le premier est qu'elle possède deux paramètres au lieu d'un ce qui permet plus de latitude pour l'adéquation de la loi théorique aux données. Son second avantage est son aptitude à représenter un phénomène à taux de panne variable (croissant ou décroissant).

Une façon simple de voir si une variable aléatoire X est bien représentable par une loi de Weibull consiste à étudier la régression linéaire du logarithme de son taux de panne par rapport au logarithme de X . La fonction les liant devrait être linéaire si l'on en croit la formule. Dans notre cas, le nuage de points $(\ln(R+1), \ln(\tau))$ obtenu est le suivant :



On constate que les points du nuage sont presque alignés au début, ensuite ils sont dispersés presque d'une manière anarchique. En réalité, cette dispersion anarchique des données des points du nuage après un certain stade s'explique par le fait suivant : le taux de panne devient extrêmement petit à partir d'un certain rang, son logarithme devient sensible à sa moindre

variation (on a en effet $\frac{d}{dx}(\ln x) = \frac{1}{x} \rightarrow +\infty$ quand $x \rightarrow 0$).

La dispersion des points du nuage après un certain stade n'implique donc pas forcément le rejet du modèle de Weibull pour la modélisation de notre série des retards.

La régression linéaire (avec Excel) de la première partie du nuage (qui correspond aux points presque alignés) nous a donné des premières valeurs pour les paramètres k et λ de la loi de Weibull à considérer.

Ensuite, en utilisant le solveur Excel, on a initialisé k et λ aux valeurs de la régression et on a cherché des valeurs voisines qui minimisent le mieux possible les deux statistiques du test χ^2 et du test de Kolmogorov-Smirnov. Les valeurs retenues au final sont :

$$k = 0.412974 \quad \text{et} \quad \lambda = 3.7711.$$

Elles donnent pour la statistique de Kolmogorov-Smirnov la valeur

$$D_n = 0.4727.$$

Comme D_n est inférieur à 1.36 (qui est la statistique de Kolmogorov-Smirnov correspondant au seuil de 5%), la loi de Weibull correspondante aux paramètres ci-dessus peut être retenue.

Par contre, le même modèle n'est pas validé par le test χ^2 . Nous pensons que ceci provient simplement de la non homogénéité des effectifs des classes.

En retenant le modèle de Weibull ci-dessus, la variable aléatoire R est considérée comme la partie entière d'une variable aléatoire continue suivant la loi de Weibull $W(k, \lambda)$. On a donc pour tout $n \in N$:

$$\begin{aligned} \Pr(R=n) &= \int_n^{n+1} \left(1 - e^{-(x/\lambda)^k} \right) dx \\ &= e^{-(n/\lambda)^k} - e^{-((n+1)/\lambda)^k}. \end{aligned}$$

C. Loi Gamma

La loi Gamma constitue aussi un bon candidat pour modéliser des phénomènes de durée de vie à taux de panne variable. Rappelons qu'il s'agit d'une loi de probabilité de support R^+ , notée $\Gamma(k, \theta)$ (avec k et θ des réels strictement positifs), de fonction de densité :

$$f(t) = \frac{t^{k-1} e^{-t/\theta}}{\Gamma(k) \theta^k},$$

Où $\Gamma(k) := \int_0^{+\infty} t^{k-1} e^{-t} dt$.

On peut montrer que le taux de panne de $\Gamma(k, \theta)$ est décroissant si $k < 1$ et est croissant si $k > 1$.

En identifiant l'espérance et la variance de notre série R par rapport à l'espérance et la variance d'une loi Gamma, on obtient des valeurs initiales pour les paramètres k et θ . Ces valeurs sont ensuite améliorées par le solveur Excel afin de minimiser la statistique Khi-deux et celle de Kolmogorov-Smirnov. Les meilleurs paramètres obtenus sont :

$$k = 0.225236 \quad \text{et} \quad \theta = 64.78542.$$

Cependant, la loi Gamma correspondant à ces paramètres s'ajuste moins bien à notre série que la loi de Weibull retenue précédemment. Les tests révèlent une adéquation aux données moins bonne.

D. Mélange de deux exponentielles

On a aussi tenté de modéliser notre série de retards R par un mélange de deux lois exponentielles, c'est-à-dire en écrivant R sous la forme :

$$R = UX_1 + (1-U)X_2,$$

avec X_1 et X_2 sont des variables aléatoires suivant des lois exponentielles $\text{Exp}(\lambda_1)$ et $\text{Exp}(\lambda_2)$ et U est une variable aléatoire discrète, prenant la valeur 1 avec une probabilité p et la valeur 0 avec la probabilité $1-p$.

La méthode du maximum de vraisemblance nous a donné comme meilleurs paramètres de ce modèle les suivants :

$$p = 0.334452, \quad \lambda_1 = 0.024929 \quad \text{et} \quad \lambda_2 = 0.828597.$$

Notons que les méthodes de minimisation des statistiques Khi-deux et de Kolmogorov-Smirnov donnent des valeurs légèrement différentes de celles-ci.

Bien que le modèle contient plus de paramètres que celui de Weibull, on constate au vu des statistiques de tests qu'il est moins bon que ce dernier pour modéliser notre série de retards.

E. Mélange de trois exponentielles

Une loi mélange de 3 exponentielles est une loi de probabilité de support R^+ et dont la fonction de densité s'écrit :

$$f(t) = p_1 \lambda_1 e^{-\lambda_1 t} + p_2 \lambda_2 e^{-\lambda_2 t} + p_3 \lambda_3 e^{-\lambda_3 t},$$

λ_1, λ_2 et λ_3 sont des réels strictement positifs et p_1, p_2 et p_3 sont des réels strictement positifs vérifiant : $p_1 + p_2 + p_3 = 1$.

Comme on le constate, la loi mélange de trois exponentielles est une loi à 5 paramètres indépendants, elle doit donc nous donner une modélisation efficace. La méthode du maximum de vraisemblance nous a donné comme meilleurs paramètres convenables à notre série de retards R , les suivants :

$$p_1 = 0.159672, \quad p_2 = 0.567589, \quad p_3 = 0.272739, \quad \lambda_1 = 0.014577, \\ \lambda_2 = 1.124861, \quad \lambda_3 = 0.099087.$$

Ce modèle est validé par le test de Kolmogorov-Smirnov, on a obtenu pour la statistique de Kolmogorov-Smirnov la valeur : $D_n = 0.1475$.

De plus, on a remarqué que les premiers effectifs théoriques (les plus nombreux) se rapprochent de façon précise des effectifs observés. Ce modèle est nettement meilleur que le modèle de Weibull sélectionné au départ. Par contre, le test du chi deux n'a pas validé le modèle. Nous remarquons donc que les probabilités $\Pr(R \leq n)$ et $\Pr(R \geq n)$ issus de la

fonction de répartition (test de Kolmogorov Smirnov) sont plus précises que les probabilités $\Pr(R = n)$ du test du Chi deux.

Conclusion

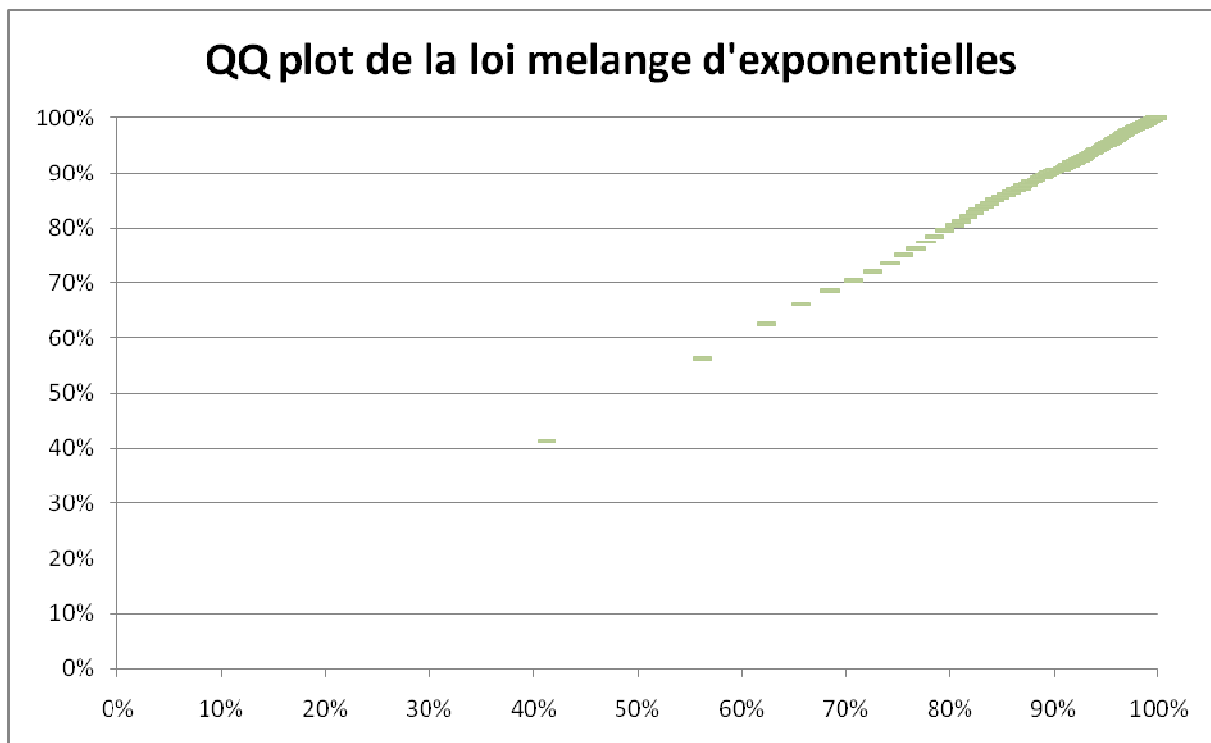
Le modèle retenu est le mélange des trois exponentielles. La loi de modélisation de R est donc la suivante :

$$\Pr(R = n) = p_1 \left(e^{-\lambda_1 n} - e^{-\lambda_1 (n+1)} \right) + p_2 \left(e^{-\lambda_2 n} - e^{-\lambda_2 (n+1)} \right) + p_3 \left(e^{-\lambda_3 n} - e^{-\lambda_3 (n+1)} \right),$$

Avec :

$$p_1 = 0.159672, \quad p_2 = 0.567589, \quad p_3 = 0.272739, \quad \lambda_1 = 0.014577, \\ \lambda_2 = 1.124861, \quad \lambda_3 = 0.099087.$$

Le QQ plot entre les valeurs observées et les valeurs théoriques confirme la bonne adéquation de la loi.



Les intervalles de confiance ont été déterminés en réitérant l'optimisation numérique sur 2000 échantillons générés par des méthodes de bootstrapping. Les intervalles sont les suivants :

$$p_1 \in [0.1546268, 0.16979004], p_2 \in [0.56206083, 0.57247602], p_3 \in [0.25782358, 0.28331237]$$

$$, \lambda_1 \in [0.01446023, 0.01522886], \lambda_2 \in [1.10678085, 1.14605416], \lambda_3 \in [0.09632647, 0.10414273]$$

F. Justification empirique du modèle retenu

En observant attentivement notre série de retards, on s'aperçoit que l'on a affaire à trois grands types de retards :

- Les retards de courte durée qui sont les plus fréquents (ils correspondent à peu près à $R \leq 7$) et sont motivés par un excès d'activité non absorbé par les équipes de gestion. Ces dernières sont conscientes du retard et saisissent les actes lorsqu'elles peuvent le faire. Les actes arrivés incomplets font l'objet d'une attente, le dossier est renvoyé chez le commercial afin d'être complété, la date de valeur de l'acte prend effet à la réception de l'intégralité de la complétude du dossier. Il s'avère qu'un certain temps (dont il est question ici) peut s'écouler avant que le gestionnaire reprenne le dossier laissé en suspens. En effet, les actes étant traités au fil de l'eau (FIFO) L'exponentielle correspondant à ce premier effet est celle dont le paramètre et la probabilité sont respectivement :

$$\lambda_2 = 1.124861 \text{ et } p_2 = 0.567589$$

On notera que pour cette exponentielle avec ce paramètre, la probabilité d'être en dessous de 7 jours de retard est très proche de 1 soit **0.9996195**.

La probabilité d'avoir quinze jours ou plus de retard de saisie avec cette loi est quasiment nulle (de l'ordre du huitième chiffre après la virgule).

Les propriétés de cette exponentielle donnent :

$$E(R_2) = 1/\lambda_2 = \sigma(R_2) = 0.88$$

Elle est donc peu dispersée autour de sa moyenne.

- Les retards de moyenne durée de fréquence intermédiaire (ils correspondent à $R \in [8, 100]$) correspondent le plus souvent à des causes identiques au premier cas mais sur des dossiers plus complexes (fiscalité étrangère, montages financiers) qui durent plus longtemps. On remarque aussi un certain nombre de dossiers pour lesquels un geste commercial est réalisé sur la date d'effet au regard de l'importance du client ou de l'investissement. Enfin, on a observé des cas de dysfonctionnement informatique sur des fonctionnalités proposées aux apporteurs d'affaires (arbitrages par internet).

L'exponentielle correspondant à ce second effet est celle dont le paramètre et la probabilité sont respectivement :

$$\lambda_3 = 0.099087 \text{ et } p_3 = 0.272739$$

Elle est plus de deux fois moins probable que la précédente.

On notera que pour cette exponentielle avec ce paramètre, la probabilité d'être en dessous de 100 jours de retard est très proche de 1 soit **0.99995**.

$$E(R_3) = 1/\lambda_3 = \sigma(R_3) = 10.092$$

- Les retards de longue durée qui sont rares (ils correspondent à peu près à $R > 100$). Ils correspondent à un rattrapage rare et ancien suite à la constatation d'une erreur de gestion (le plus souvent mauvais paramétrage de taux dérogé suite à négociation commerciale) ou un problème informatique non décelé. L'erreur impose la reprise de l'historique des actes de gestion.

L'exponentielle correspondant à ce troisième effet est celle dont le paramètre et la probabilité sont respectivement :

$$\lambda_1 = 0.014577 \text{ et } p_1 = 0.159672$$

Elle est de deux fois moins probable que la précédente.

On notera que pour cette exponentielle avec ce paramètre, la probabilité d'être au dessous de 100 jours de retard est supérieure à **0.25**.

$$E(R_1) = 1/\lambda_1 = \sigma(R_1) = 68.6$$

Ce troisième effet est beaucoup plus dispersé autour de sa moyenne.

L'espérance mathématique de la loi mélange du retard se déduit des résultats classiques et connus des lois exponentielles et de la linéarité de l'espérance, soit :

$$E(R) = (p_2 / \lambda_2) + (p_3 / \lambda_3) + (p_1 / \lambda_1) = 0.5 + 2.75 + 10.95 = 14.2$$

On insiste par ailleurs sur le fait que la dispersion des retards n'est pas la même selon les trois effets puisque pour une exponentielle, l'écart type est inversement proportionnel au paramètre de la loi et est égale à l'espérance.

Le modèle retenu (mélange de trois lois exponentielles) conduit à reconnaître ces trois types de retards.

V. MODELISATION DES RCE 2008 & 2007

A. Méthodologie utilisée

Nous rappelons les formules de calcul du Résultat de couverture sur un flux, nous allons déterminer alors la distribution des RCE par des méthodes de monte Carlo sous l'hypothèse des flux 2008 avec des échantillons de retards et de cours simulés à partir des lois déterminées ci-haut. Nous ferons de même avec la distribution des flux constatés en 2007.

Nous chercherons alors à montrer qu'en faisant abstraction de l'effet « flux » (en standardisant en montants et rapports entre les cash flows positifs et négatifs les deux séquences de flux 2007 et 2008), les deux distributions de RCE obtenues proviennent statistiquement de la même loi. La distribution des flux est très différente entre les deux années (2008 année de crise financière vs 2007 année de croissance boursière). Nous modéliserons alors dans la partie VI, la distribution des flux demandés par les assurés.

B. Rappel des formules de calcul

Les actes classiques saisis sans retard prennent effet le lendemain. Ils sont estimés à la dernière valeur liquidative connue au moment de la saisie (dans le cas d'un support en cotation quotidienne, c'est la VL du deuxième jour ouvré précédent la saisie).

Sa valeur est donc naturellement égale à :

$$\text{Ordre d'achat} = \frac{\text{Montant}}{\text{VL saisie}} \quad (\text{En nombre de parts}).$$

Où *Montant* désigne le montant algébrique investi par le client et *VL saisie* est la valeur liquidative connue à la date de saisie (i.e celle de deux jours ouvrés avant)

Le flux représente le nombre de parts qu'on aurait du acheter à la date d'effet, sa valeur est donc égale à :

$$\text{Flux} = \frac{\text{Montant}}{\text{VL effet}}$$

Où *VL effet* est la valeur liquidative à la date d'effet.

La date d'ajustement est la date à laquelle les flux sont ajustés. Si *M* désigne le mois de la date de saisie, la date d'ajustement serait toujours le 25 du mois (*M* + 2) excepté le cas où cette dernière date tombe un week-end ou un jour férié. Dans ce cas, la date d'ajustement sera le premier jour ouvré venant après cette date.

A la date d'ajustement, on adosse l'actif en parts sur le passif en parts par l'intermédiaire d'un flux d'ajustement calculé en réalité à partir d'un écart sur un stock de fin de mois, mais que l'on considèrera ici comme réalisé sur chaque flux algébrique. Le nombre de parts correspondant au flux de régularisation est donné par :

$$\text{Regul} = \text{Flux} - \text{Ordre d'achat}$$

Ces parts de trop (ou de moins) ont été achetées (ou vendues) au prix de VL_{effet} la part alors qu'à la date d'ajustement, chaque part vaudra VL_{ajust} . Notre gain (resp. perte) est donc égale à :

$$REGUL = (VL_{effet} - VL_{ajust})$$

Cette dernière valeur s'appelle le *Résultat de couverture* et se note par RCE. On retient donc la formule principale pour le résultat de couverture relatif à un acte de gestion classique:

$$\begin{aligned} RCE &= Regul \times (VL_{effet} - VL_{ajust}) \\ &= \left(\frac{\text{Montant}}{VL_{effet}} - \frac{\text{Montant}}{VL_{saisie}} \right) \times (VL_{effet} - VL_{ajust}) \end{aligned}$$

C. Formule du RCE pour les actes de gestion rétroactifs

Les ordres d'achats des actes rétroactifs pour leur part sont calculés à la valeur d'effet au passif car au moment du passage d'ordre, on connaît déjà la valeur liquidative au passif et donc aucune estimation de nombre de parts n'est nécessaire. Un raisonnement similaire à celui de ci-dessus aboutit à la formule suivante du RCE relative à un acte de gestion rétroactif :

$$RCE = \frac{\text{Montant}}{VL_{effet}} (VL_{effet} - VL_{actif}).$$

Où VL_{actif} est la valeur liquidative du lendemain de la date de saisie, valeur liquidative à laquelle sera réalisée l'opération chez le dépositaire.

D. L'approche monte Carlo

Nous venons de démontrer que le RCE (résultat de couverture économique) sur un flux est complètement défini par quatre paramètres : le montant financier du flux, la date de saisie, le retard de saisie, la séquence des valeurs liquidatives.

Nous avons modélisé les retards par une loi mélange d'exponentielles.

Nous avons représenté les trajectoires financières de nos unités de compte par un modèle RSLN2 issu de l'historique du CAC40.

Nous disposons de la séquence des flux saisis au passif sur l'année 2008.

Nous allons chercher à approcher la distribution de nos résultats de couverture économique annuels en utilisant les méthodes stochastiques dites de Monte Carlo.

La méthode de simulation de Monte Carlo consiste à envisager un grand nombre de trajectoires du RCE en simulant pour chaque flux, un retard de saisie et un scénario financier.

Nous utilisons alors les nombres obtenus pour calculer un RCE sur ce flux.

La somme algébrique des RCE calculés individuellement sur chaque flux et pour chaque date de saisie, doit nous donner un montant de RCE annuel.

Nous pouvons donc obtenir pour un grand nombre de scénarios (ce nombre sera choisi suffisamment élevé de telles sortes que la convergence vers une distribution théorique soit atteinte) une distribution des RCE annuels théoriques de laquelle nous pouvons estimer un certain nombre d'indicateurs sur le résultat de couverture économique.

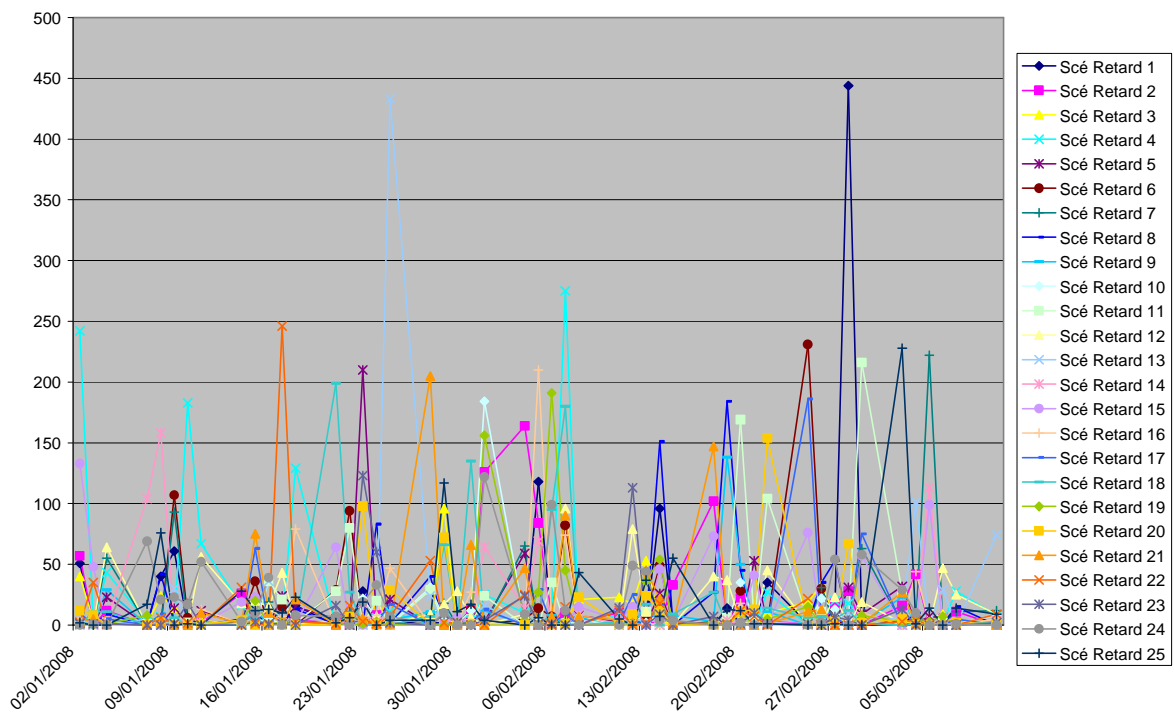
1. Comment simuler un retard et un scénario financier ?

Pour simuler différentes réalisations de lois de probabilité (voir Article de PLANCHET F. et JACQUEMIN J. sur *L'utilisation de méthodes de simulation en assurance* Bulletin Français d'Actuariat (numéro 11) – Volume 6 – Juin / Décembre 2003), nous utilisons la génération de nombres aléatoires de loi uniforme et les méthodes type « inversion de la fonction de répartition ».

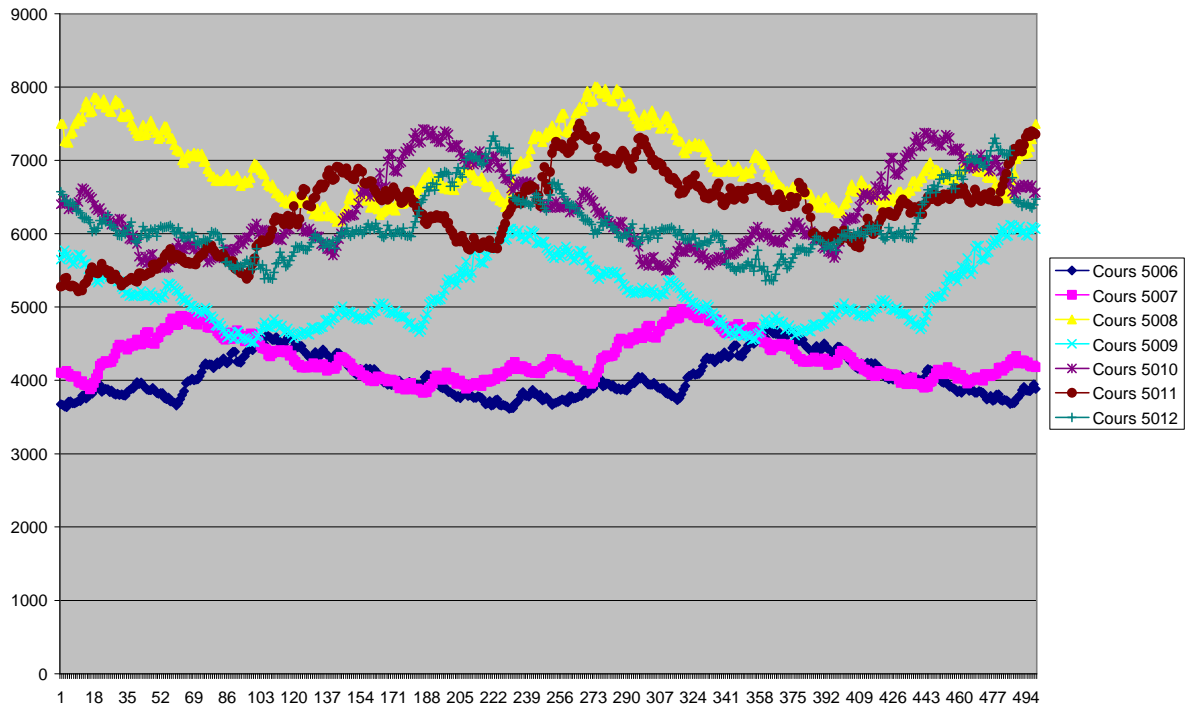
Ainsi pour les retards, après avoir généré un nombre entre 0 et 1, nous allons chercher le premier entier n tel que la fonction de répartition de la loi mélange d'exponentielle appliquée en n soit supérieur au nombre généré.

Des techniques analogues sont mises en œuvre pour simuler un scénario RSLN2.

Extrait d'une simulation de retards (en jours par date de saisie)

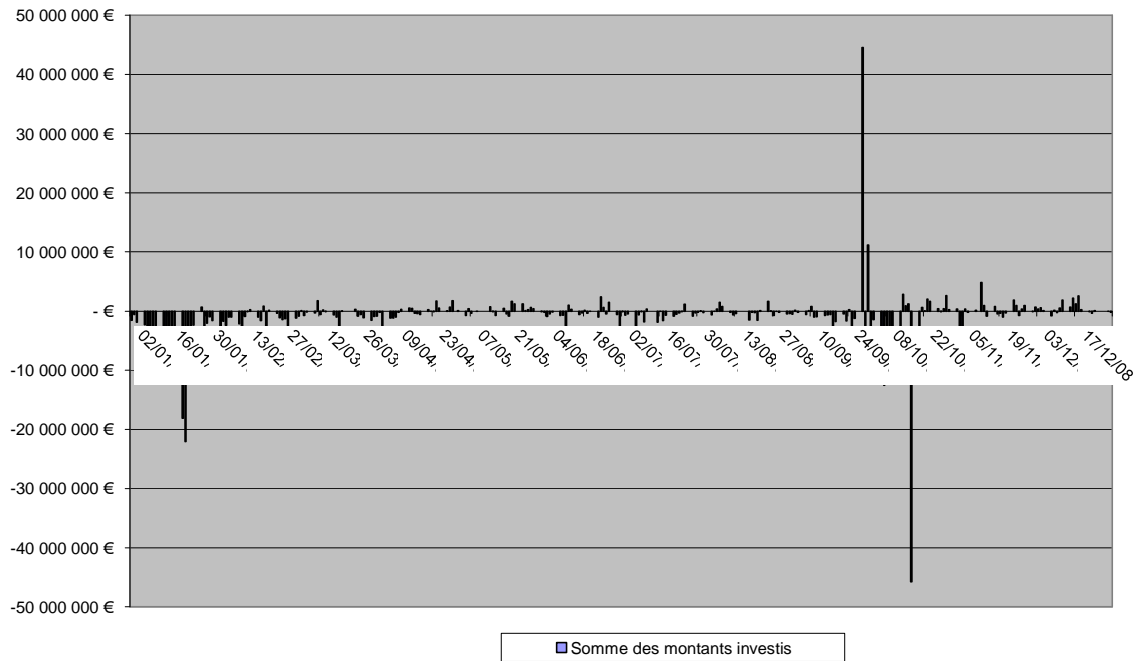


Extrait d'une simulation de cours (avec une valeur de CAC égale au cours du 02/01/08 pour toutes les simulations



2. Structure des flux 2008

Cash flows 2008 par date de saisie



Les cash flows 2008 se caractérisent par une majorité de montants négatifs (392 M€ et 171 flux sur 251 soit 2/3 des flux en nombre et 78% des cash-flows en valeur absolu)

On constate deux valeurs extrêmes correspondant à la crise financière : + 44M€ le 29/09/08 et -45 M€ le 17/10/08.

Le premier chiffre correspond à quelques arbitrages sur un petit nombre de contrats à encours élevé en sortie d'un fonds en euro et en entrée sur un fond majoritairement obligataire.

Le second est directement imputable à la « panique » des premiers jours de la crise financière 2008.

3. Distribution du RCE 2008

Nous avons simulé 10 000 couples de scénarios (retard, cours) puis de nouveau 5000 autres scénarios et encore 5000 distincts des précédents. IL faut noter que ces opérations se sont révélées très chronophages. Par ailleurs, le calcul du RCE flux par flux sur les 400 000 données de l'année 2008 s'est révélé impossible pour une question de taille de base de données et de puissance de calcul même avec le nombre minimal à nos yeux de 10 000 scénarios de retards et 10 000 scénarios de RSLN2.

Nous avons donc agrégé nos données par date de saisie et obtenu 254 cash flows correspondant aux montants saisis pour chacun des jours de saisie 2008 (jours ouvrés et non fériés). Nous avons pu alors procéder aux simulations.

L'année 2008 étant une année de crise financière caractérisée par une majorité de cash flows négatifs. L'inconvénient est que le retard simulé par la loi s'applique à l'ensemble des cash flows de la journée pour chacune de nos itérations. Tout se passe en réalité comme si les cash flows de la journée étaient saisis sur un seul acte avec le retard simulé. Nous savons que plus le retard est grand, plus il est improbable, nous ne pouvons donc mettre sur un pied d'égalité pour un montant donné, les résultats de couverture simulés avec un retard faible et ceux générés par un retard de saisie conséquent, Nous risquons donc en fait d'aggraver la queue de distribution que nous cherchons à maîtriser. Nous devons donc pondérer les résultats de couverture simulés par les fréquences observées des retards de saisie sur une journée.

a. Estimation des fréquences de saisie

A partir des fréquences observées de saisie de l'année 2008, nous avons calculé les moyennes observées par jour pour chaque date de saisie 2008.

Nous avons discrétisé l'espace des retards possibles en intervalles disjoints tels que nous puissions déterminer une fréquence observée sur chaque intervalle (certains retards n'ont évidemment jamais pu être observés). Cela revient donc à estimer les probabilités d'occurrence pour chacun des événements possibles d'une loi multinomiale.

L'estimateur du maximum de vraisemblance de ces probabilités est bien l'ensemble des proportions moyennes de retards observés. Si un retard n'est pas observé sur une tranche donnée, la proportion observée est bien évidemment nulle.

Les retards nuls ou faibles (une à deux journées) sont suffisamment observés pour faire l'objet de classe à part entière.

A partir de trois jours, nous commençons par faire des classes de deux journées puis progressivement nous élargissons la taille de la classe à mesure que le besoin s'en fait sentir.

Les classes sont de taille quatre à partir du cinquante et unième jour de retard.
 Elles sont de taille dix à partir du cent quatrième jour de retard.
 La dernière classe correspond aux retards supérieurs ou égaux à 350 jours.

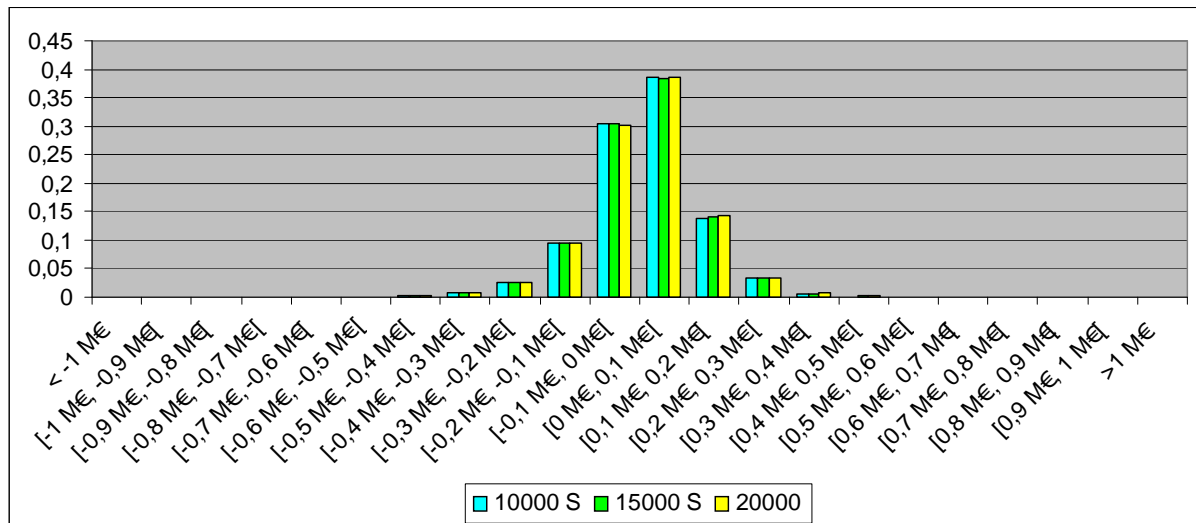
Nous avons donc obtenu les fréquences suivantes de saisie sur une journée :

| Tranche de Retards corrigés | Fréquence |
|------------------------------------|------------------|
| 0 | 37,81% |
| 1 | 17,05% |
| 2 | 4,97% |
| [3, 4] | 5,20% |
| [5, 6] | 3,84% |
| [7, 8] | 3,61% |
| [9, 10] | 1,92% |
| [11, 12] | 1,62% |
| [13, 14] | 1,36% |
| [15, 17] | 2,35% |
| [18, 20] | 1,56% |
| [21, 23] | 1,44% |
| [24, 26] | 0,92% |
| [27, 30] | 1,53% |
| [31, 34] | 1,31% |
| [35, 38] | 1,09% |
| [39, 42] | 0,94% |
| [43, 46] | 0,85% |
| [47, 50] | 0,68% |
| [51, 55] | 0,84% |
| [56, 60] | 0,72% |
| [61, 65] | 0,60% |
| [66, 71] | 0,55% |
| [72, 77] | 0,45% |
| [78, 83] | 0,44% |
| [84, 89] | 0,39% |
| [90, 96] | 0,51% |
| [97, 103] | 0,33% |
| [104, 114] | 0,57% |
| [115, 125] | 0,51% |
| [126, 136] | 0,37% |
| [137, 147] | 0,30% |
| [148, 158] | 0,24% |
| [159, 169] | 0,23% |
| [170, 190] | 0,46% |

| | |
|------------|-------|
| [191, 220] | 0,84% |
| [221, 250] | 0,50% |
| [251, 350] | 0,52% |
| [350, +∞ [| 0,57% |

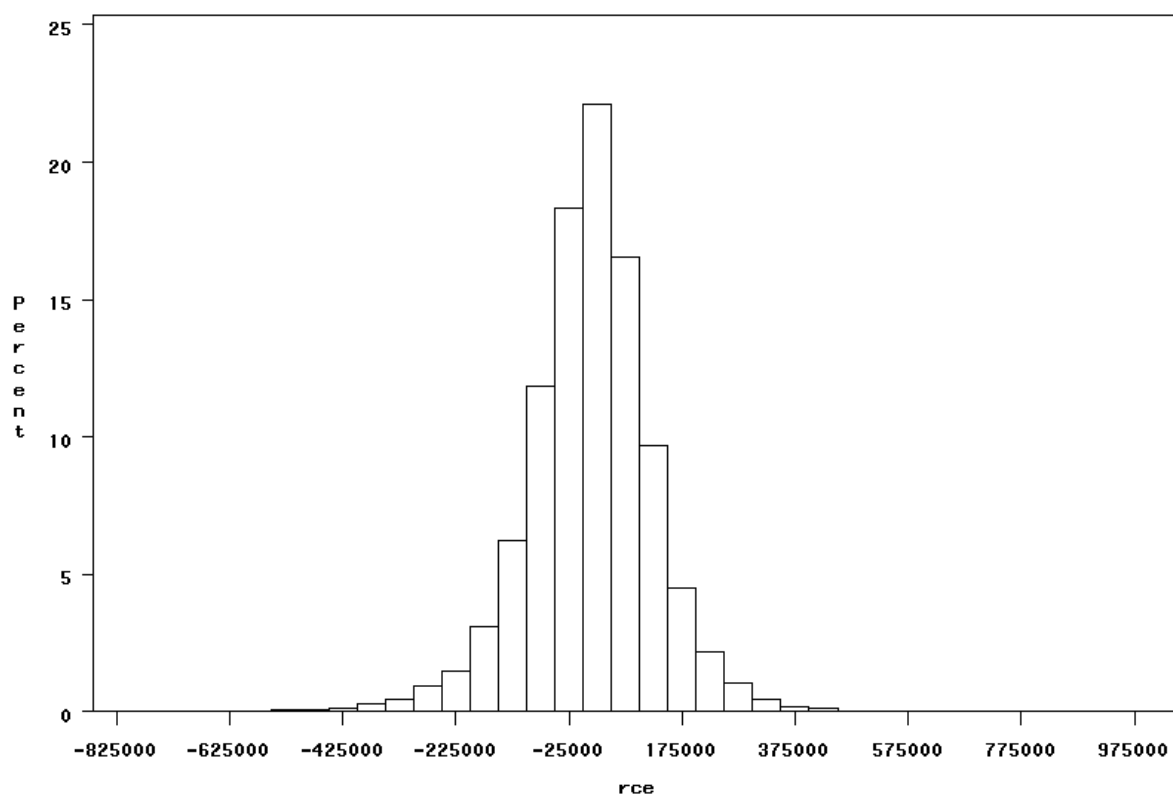
On a estimé l'espérance et les quartiles du RCE 2008 en utilisant les 10000 premiers scénarios puis les 15000 et enfin tous les 20000 scénarios. La comparaison des résultats obtenus suggère que la suite des distributions de nos scénarios est assez proche de sa limite, c'est-à-dire de la distribution « réelle » du RCE.

b. Distribution par tranche de RCE observés



La même distribution vue avec des tranches d'histogrammes différentes :

distribution des rce 2008



Les indicateurs empiriques observés du RCE sont les suivants :

| Nb de scénarios | Min | Moyenne | Quartile 1 | Médiane | Quartile 3 | Max |
|-----------------|-------------|----------|------------|----------|------------|-------------|
| 10 000 | - 835 172 € | 11 459 € | - 50 459 € | 15 066 € | 76 747 € | 849 925 € |
| 15 000 | - 835 172 € | 11 868 € | - 49 503 € | 14 674 € | 77 382 € | 849 925 € |
| 20 000 | - 835 172 € | 12 281 € | - 49 115 € | 15 885 € | 77 677 € | 1 042 342 € |

Par ailleurs on calcule l'écart type empirique sans biais suivant :

| Nb de scénarios | Ecart type |
|-----------------|------------|
| 10 000 | 113 192 € |
| 15 000 | 113 575 € |
| 20 000 | 113 436 € |

La distribution semble symétrique (la médiane est proche de la moyenne) et la moyenne est proche de zéro. Les risques de perte ou de gain sont équidistribués sur une étendue de pratiquement 2 M€.

Ceci nous est confirmé par le calcul du coefficient d'asymétrie (Skewness) qui est proche de 0 soit -0.21 (négatif donc légèrement allongé vers la gauche c'est-à-dire vers les RCE positifs). En revanche, le coefficient d'aplatissement qui mesure le degré de concentration des observations dans les queues (Kurtosis normalisé) est de 3.019 (différent de 0).

Les queues de distribution ont plus de « poids » que pour une distribution normale. (Un kurtosis positif indique que les queues comptent plus d'observations que dans une distribution gaussienne et réciproquement).

Différents tests de student et des rangs signés ne nous permettent pas de conclure à la nullité de la moyenne (student) ou de la médiane (rangs signés) comme le montrent les résultats suivants obtenus avec la proc univariate sous SAS.

| Tests de position : Mu0=0 | | | | |
|---------------------------|---|---------------|----------|-------------------|
| Test | | -Statistique- | | -----p Value----- |
| t de Student | t | 15.31111 | Pr > t | <.0001 |
| Signe | M | 1390 | Pr >= M | <.0001 |
| Rangs signés | S | 16013963 | Pr >= S | <.0001 |

Par contre l'hypothèse de normalité de la loi n'est pas non plus confirmée par les tests disponibles (sous SAS univariate) de Kolmogorov- Smirnov, Cramer-Von Mises, Anderson-Darling comme l'attestent les p values calculés rejetant la normalité au niveau de 5%.

| Tests de normalité | | | | |
|--------------------|------|----------------|-----------|-------------------|
| Test | | -Statistique-- | | -----p Value----- |
| Kolmogorov-Smirnov | D | 0.044937 | Pr > D | <0.0100 |
| Cramer-Von Mises | W-Sq | 16.1991 | Pr > W-Sq | <0.0050 |
| Anderson-Darling | A-Sq | 100.0573 | Pr > A-Sq | <0.0050 |

Sous l'hypothèse des cash flows observés en 2008, le résultat de couverture incluant nos risques (retard de saisie volontaires ou non, erreurs de saisie, dysfonctionnements des systèmes d'information, gestes commerciaux) est en moyenne légèrement positif pour la compagnie. L'assureur peut supporter cette charge qui se solde d'ailleurs par un résultat légèrement positif en moyenne. On a déjà vu (voir 1ere partie) que le résultat de couverture empirique observé réellement sur 2008 est de 1.4M€ et de 143 K€ abstraction faite des valeurs extrêmes. Par ailleurs, le montant des provisions mathématiques en UC de notre portefeuille s'élève à 1.8 milliard d'euros. Le RCE rapporté aux PM oscille dans l'intervalle [-0.05%,0.06%]

Il n'a donc pas besoin au vu de nos résultats d'engager des ressources supplémentaires à savoir provisionner, augmenter les ressources (frais d'entrée, frais de gestion, frais d'arbitrage...).

La seule contrainte qu'il nous faut analyser est la capacité de faire face aux scénarios les plus défavorables et les plus coûteux (queues de distribution que nous avons vus relativement importantes).

Nous allons devoir étudier notre risque au travers de quelques indicateurs statistiques de queue de distribution.

Remarque importante

Nous devons cependant affranchir notre distribution de l'effet des cash flows.

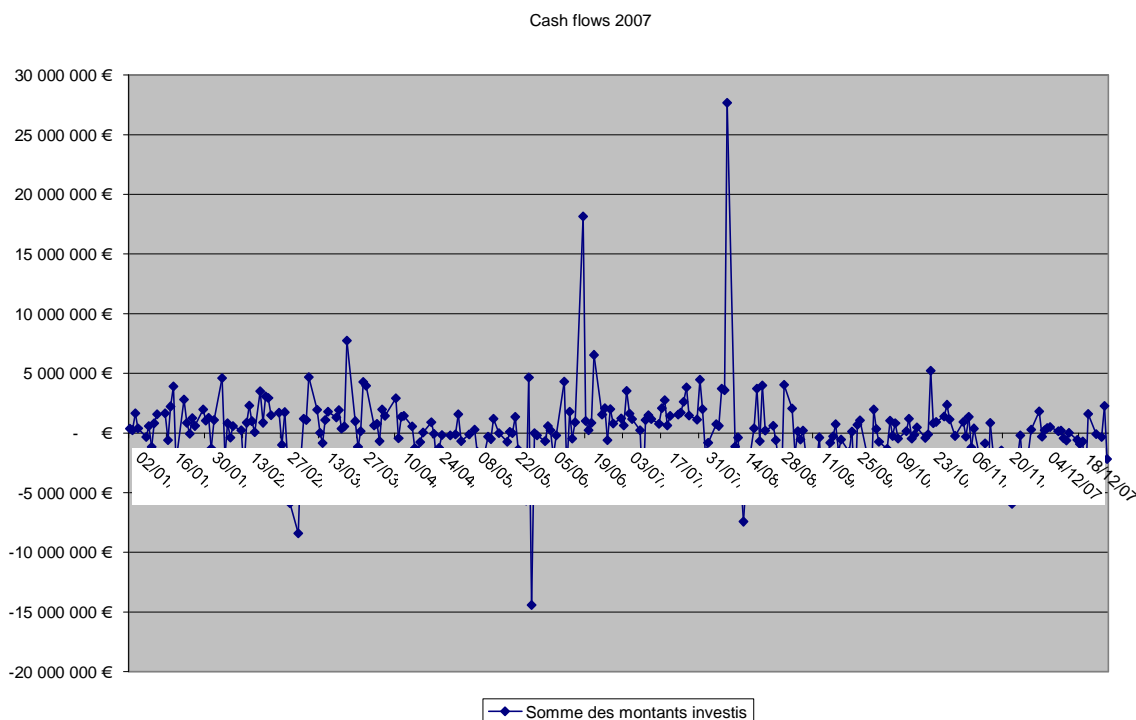
Pour cela, nous allons évaluer la distribution 2007 et établir une distribution standardisée sur la base d'un volume de cash flows annuels en valeur absolue de 1 Milliard d'euros.

En effet, Nous ne pouvons pas prendre les cash flows algébriques car il y aurait sous évaluation des risques (par compensation des montants) et par ailleurs nous souhaitons standardiser une mesure de risque sur un volume de travail au service client pour lequel les flux financiers ne s'annihilent pas.

4. Etude de la distribution réalisée à partir des cash flows 2007

Pour étudier l'impact de l'effet cash flow, nous réitérons notre étude avec la séquence des cash flows 2007.

a. Structure des flux 2007



Les cash flows 2007 se caractérisent par une majorité de montants positifs (121 M€ et 147 flux sur 253 soit 60% des flux en nombre et 58% des cash-flows en valeur absolue).

L'année 2007 est donc très différente de 2008 par sa structure de cash flows. Elle se solde par un investissement global sur l'année de 73 M€ à opposer au solde négatif de 2008 (-280 M€).

b. Distribution par tranche de RCE observés

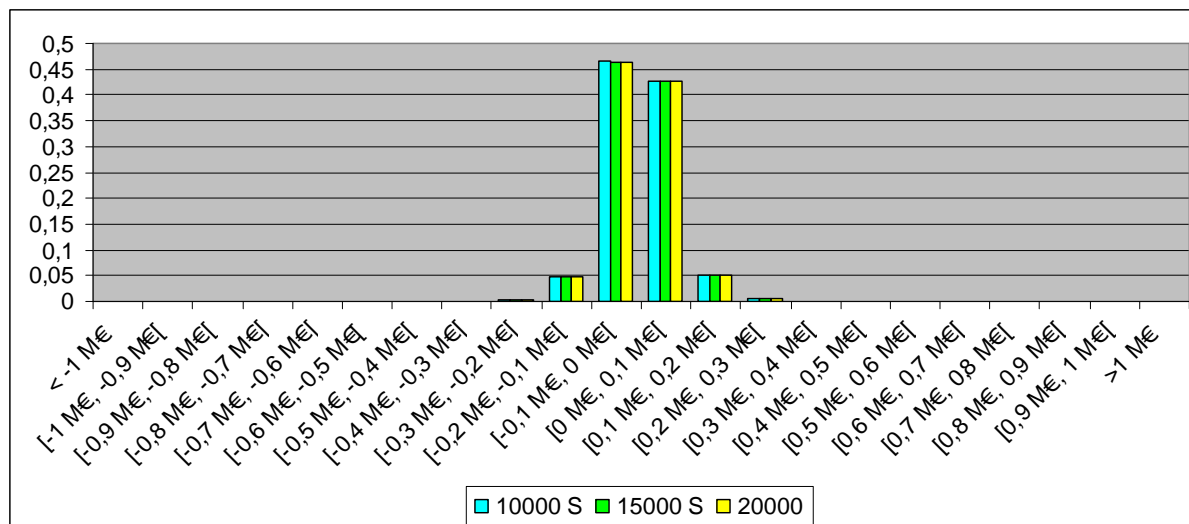
Les simulations avec la structure des cash flows 2007 ont donné les résultats suivants :

| Nb de scénarios | Min | Moyenne | Quartile 1 | Médiane | Quartile 3 | Max |
|-----------------|------------|-----------|------------|-----------|------------|-----------|
| 10 000 | -358 709 € | - 1 150 € | - 38 694 € | - 2 433 € | 35 665 € | 418 975 € |
| 15 000 | -526 670 € | - 1 376 € | - 38 991 € | - 2 143 € | 35 546 € | 418 975 € |
| 20 000 | -526 670 € | - 1 632 € | - 39 030 € | - 2 344 € | 35 196 € | 431 479 € |

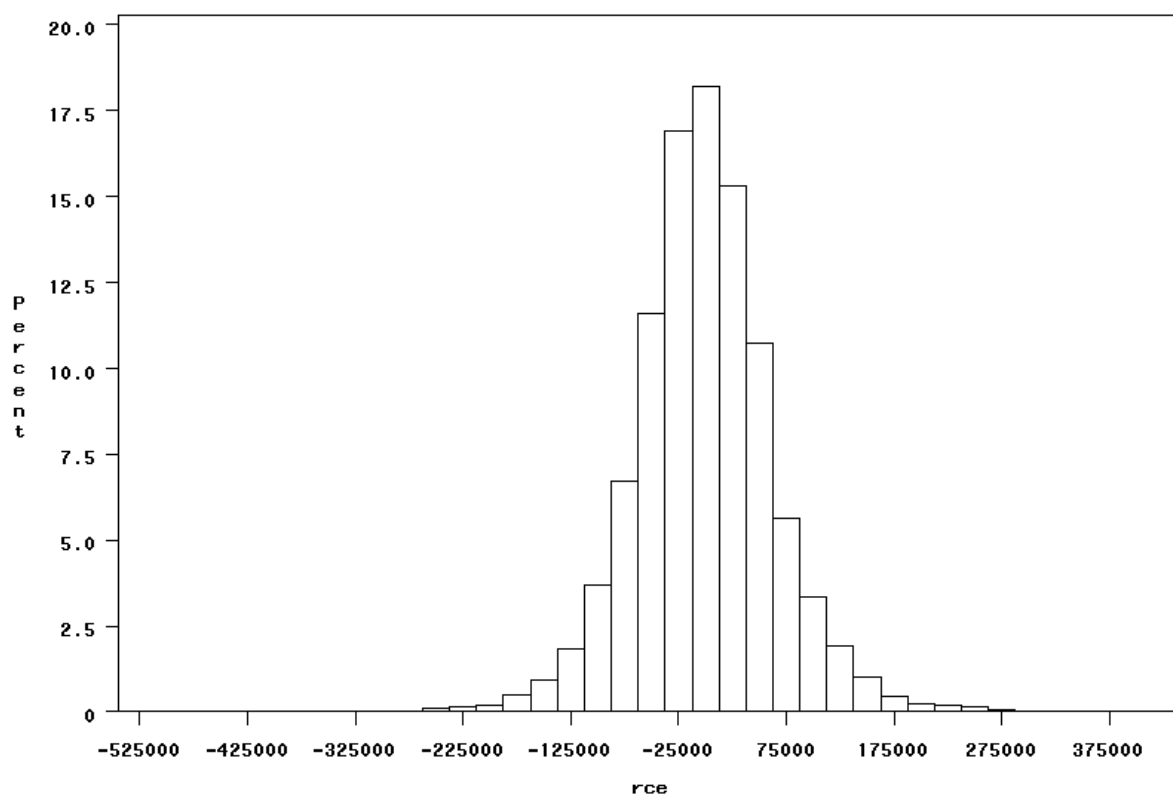
Les écarts types observés sont les suivants :

| Nb de scénarios | Ecart type |
|-----------------|------------|
| 10 000 | 64 789 € |
| 15 000 | 64 794 € |
| 20 000 | 64 766 € |

Distribution par tranche de RCE observés



distribution des rce 2007



Nous constatons que :

- La méthode de monte Carlo converge.
- La moyenne est proche de 0 mais légèrement négative.
- La distribution est symétrique (skewness de 0.0854).
- L'écart type et les queues de distribution sont plus faibles.
- La conclusion sur l'absence de ressources supplémentaires constatée sur les résultats 2008 s'applique également ici.
- L'observation des autres quantiles montre une distribution 2007 moins dispersée qu'en 2008.
- La courbe 2007 n'est pas une simple translation ou relation homothétique de la courbe 2008. Nous en déduisons que l'effet de la structure des cash flows (négatifs ou positifs) dans le calcul de nos RCE n'est pas neutre.

Nous allons le démontrer en ramenant les distributions de cash flows 2007 et 2008 à des éléments comparables entre eux.

5. Comparaison des distributions 2007 et 2008 standardisées

a. Standardisation

Nous supposons que les montants et le sens des cash flows ont un effet non neutre sur le résultat de couverture.

En effet, nous avons vu que le sens du cash flow (flux positif ou négatif) est partie intégrante de la détermination du sens du RCE (voir plus haut). Par ailleurs pour les montants, c'est plus qu'évident. Nous devons donc ramener l'un et l'autre à des dimensions comparables pour apprécier la justesse de notre distribution.

Pour cela nous allons calculer la somme des valeurs absolues des cash flows annuels et diviser l'ensemble de nos RCE par ce montant pour l'année 2008 (pour obtenir les RCE annuels pour un euro de cash flow).

En effet, nous cherchons à évaluer notre RCE pour un montant standard de flux au service client, chose qui ne peut se faire en valeurs algébriques en raison des compensations entre montants positifs et montants négatifs.

Pour 2007, nous allons au préalable diviser les cash flows positifs par 0.2045, coefficient calculé avec le solveur d'Excel permettant d'obtenir le même rapport entre les cash flows positifs et négatifs que 2008 soit 22% de cash flows positifs (rappel 2007 constaté 58%). Nous divisons alors l'ensemble des 20 000 RCE par la somme des valeurs absolues des nouveaux cash flows obtenus et nous standardisons pour l'ensemble des deux années en multipliant tous les RCE obtenus par un montant de 1 Mrds d'euros de flux saisis au service client.

b. Distributions 2007 & 2008 standardisées

Les quantiles obtenus pour **2008** sont les suivants :

| Nb de scénarios | Min | Moyenne | Quartile 1 | Médiane | Quartile 3 | Max |
|-----------------|--------------|----------|-------------|----------|------------|-------------|
| 10 000 | -1 670 344 € | 22 917 € | - 100 917 € | 30 131 € | 153 493 € | 1 699 849 € |
| 15 000 | -1 670 344 € | 23 735 € | - 99 006 € | 29 348 € | 154 763 € | 1 699 849 € |
| 20 000 | -1 670 344 € | 24 562 € | - 98 231 € | 31 770 € | 155 354 € | 2 084 683 € |

| Nb de scénarios | Ecart type |
|-----------------|------------|
| 10 000 | 226 384 € |
| 15 000 | 227 149 € |
| 20 000 | 226 872 € |

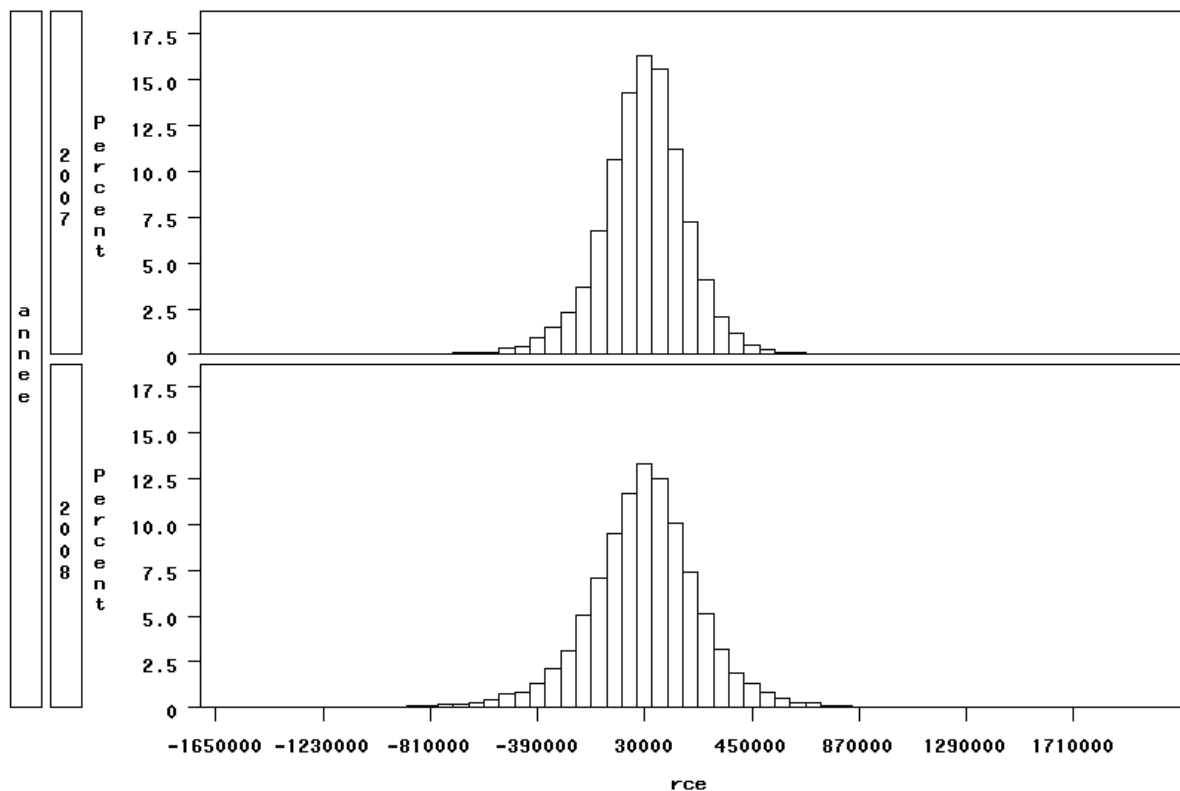
Pour 2007, les quantiles obtenus sont les suivants

| Nb de scénarios | Min | Moyenne | Quartile 1 | Médiane | Quartile 3 | Max |
|-----------------|--------------|----------|------------|----------|------------|-----------|
| 10 000 | -1 125 050 € | 23 924 € | - 72 923 € | 30 544 € | 128 209 € | 996 656 € |
| 15 000 | -1 125 050 € | 25 521 € | - 70 454 € | 31 566 € | 128 004 € | 996 656 € |
| 20 000 | -1 125 050 € | 25 723 € | - 69 983 € | 32 234 € | 127 758 € | 996 656 € |

| Nb de scénarios | Ecart type |
|-----------------|------------|
| 10 000 | 171 222 € |
| 15 000 | 169 795 € |
| 20 000 | 170 308 € |

Les indicateurs de moyenne sont assez proches. Conformément à ce que l'on a déjà vu, la distribution 2007 est légèrement moins dispersée. Ce que l'on peut visuellement observer sur la représentation graphique suivante :

distribution des rce 2007 vs 2008



Nous devons valider l'hypothèse selon laquelle les deux distributions proviennent de la même loi. Pour cela nous utiliserons deux tests non paramétriques (pas d'hypothèse faite sur la distribution des variables).

c. Le test de Wilcoxon

Le principe du test de Wilcoxon-Mann-Whitney consiste à ranger dans l'ordre croissant l'ensemble des observations mélangées des deux échantillons, à leur affecter un rang et à calculer séparément la somme des rangs des observations provenant de chacun des deux échantillons (notées QX et QY). Sous l'hypothèse H0 d'égalité des distributions issues des deux échantillons, la statistique

$$U = Q_x - \frac{(n+1)*n}{2} \text{ ou } n \text{ est la taille du plus petit échantillon suit une loi normale.}$$

Si la statistique observée est plus petite que 0 (elle vaut 0 sous H0 lorsqu'elle est centrée réduite)

La procédure **NPAR1WAY** de **SAS** donne les p values suivantes :

$$P(U \leq u) \ \& \ P(|U| > |u|)$$

L'utilisation de rangs au lieu de valeurs numériques complètes n'est pas seulement une question de commodité occasionnelle. Elle a aussi des conséquences importantes et utiles car les rangs ne sont pas affectés par un changement d'échelle sur x , même si ce changement d'échelle n'est pas uniforme sur le domaine de x . En fait, les rangs ne changent pas (sont "invariants") sous toute transformation **monotone** de x . Ces tests ne reposent alors sur aucune hypothèse concernant les distributions sous-jacentes aux données et sont donc plus robustes que les tests paramétriques.

Les 2 p values obtenues sont respectivement de 49% et 98% ce qui nous incline à accepter l'identité des distributions de façon confortable au regard du risque de 1ere espèce.

On notera au passage la forte proximité des scores observés avec ceux attendus sous H0.

Test de Wilcoxon & Kruskal-Wallis

The NPAR1WAY Procedure

Wilcoxon Scores (Rank Sums) pour la variable rce
Classée par variable annee

| annee | Nb | Somme des scores | Attendue sous H0 | Écart-type sous H0 | Score moyen |
|-------|-------|------------------|------------------|--------------------|-------------|
| 2008 | 20000 | 399992832 | 400010000 | 1154714.97 | 19999.6416 |
| 2007 | 20000 | 400027168 | 400010000 | 1154714.97 | 20001.3584 |

Wilcoxon Two-Sample Test

| | |
|---------------------------|----------------|
| Statistic (S) | 399992832.0000 |
| Normal Approximation | |
| Z | -0.0149 |
| One-Sided Pr < Z | 0.4941 |
| Two-Sided Pr > Z | 0.9881 |
| t Approximation | |
| One-Sided Pr < Z | 0.4941 |
| Two-Sided Pr > Z | 0.9881 |
| Exact Test | |
| One-Sided Pr <= S | . |
| Two-Sided Pr >= S - Mean | . |

Z includes a continuity correction of 0.5.

d. Le test de Kruskal-Wallis

C'est une généralisation du test précédent. La statistique du test de Kruskal-Wallis est construite à partir des **moyennes** des rangs des observations dans les différents échantillons que l'on compare entre elles et qui doivent être égales sous l'hypothèse d'identité des distributions (H0)

Le niveau de la pvalue est 98%. On accepte donc H0.

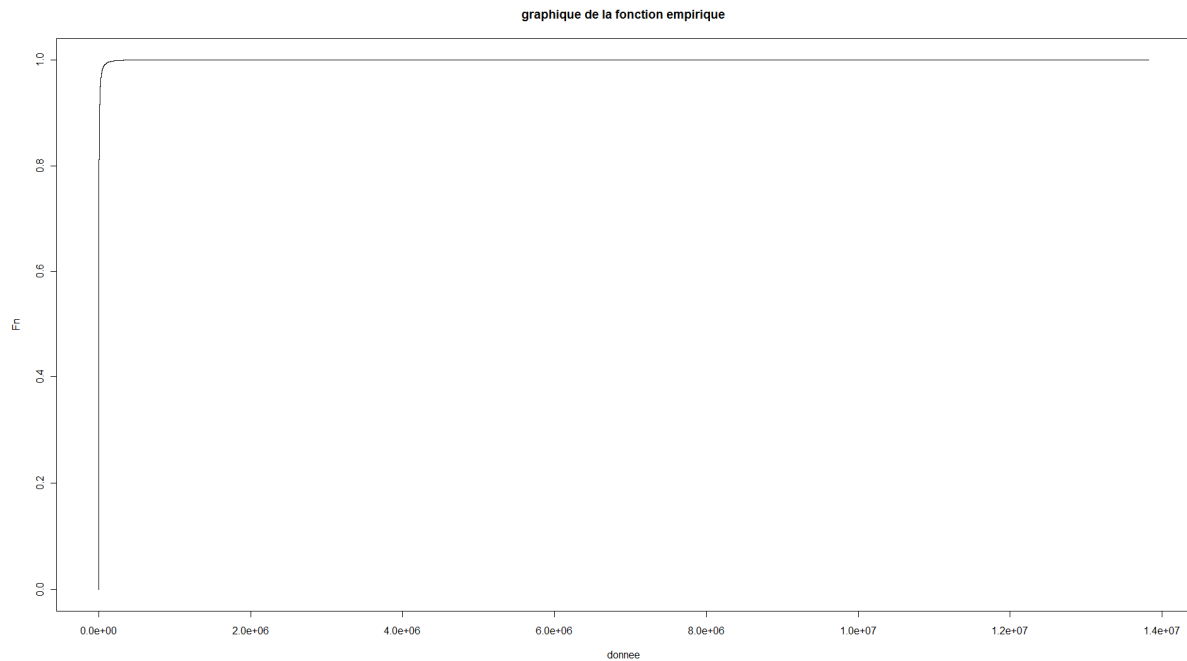
| | |
|---------------------|--------|
| Kruskal-Wallis Test | |
| Chi-Square | 0.0002 |
| DF | 1 |
| Pr > Chi-Square | 0.9881 |

Nous venons de démontrer que la distribution des RCE annuels est invariante lorsque le rapport entre les cash flows positifs et négatifs est stable et lorsque les volumes des montants globaux en valeur absolue sont ramenés à une échelle comparable. Nous allons donc tenter de chercher une modélisation théorique à la distribution des montants.

Nous avons constaté qu'il n'y a pas de distinction significative entre les montants de chiffre d'affaires et les prestations. Nous allons donc chercher à modéliser l'ensemble des flux par une loi appropriée afin de pouvoir générer des scénarios que l'on incorporera dans nos simulations monte Carlo. Nous utiliserons pour cela l'ensemble de la distribution de la valeur absolue des montants. Nous allons donc tester différentes lois usuelles connues après avoir analysé la distribution de valeurs absolues des montants.

VI. MODELISATION DES FLUX FINANCIERS DES ASSURES

A. Fonction de répartition empirique

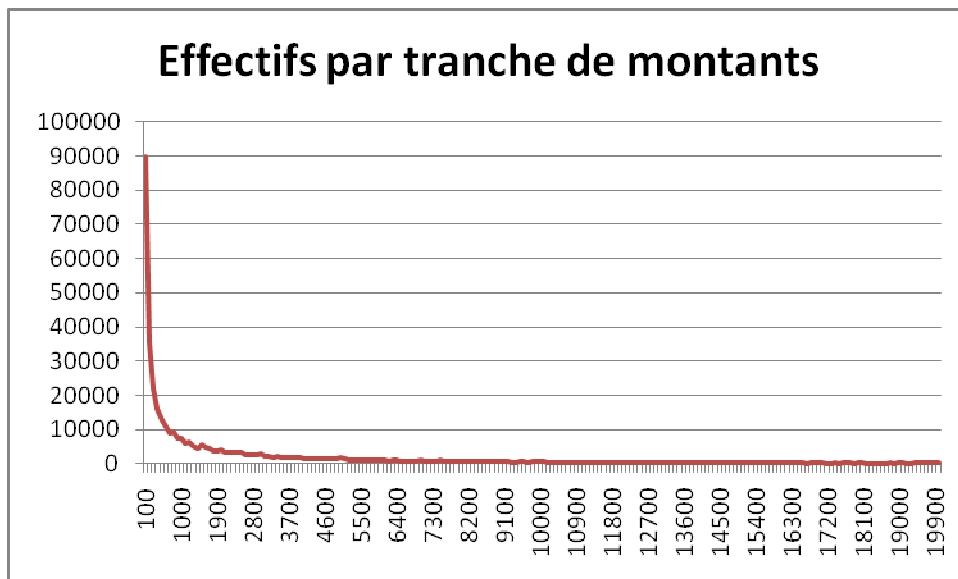


Les principaux quantiles observés sont les suivants :

| quantile | Min | Moyenne | Quartile 1 | Médiane | Quartile 3 | Max |
|-----------------|------------|----------------|-------------------|----------------|-------------------|--------------|
| Montant | 30 € | 6 881 € | 236 € | 1 190 € | 4 719 € | 13 820 000 € |

Nous constatons une distribution très dense sur des petits montants avec des valeurs très élevées statistiquement très peu observées. Cela doit nous faire penser à la loi de Pareto et généralement à l'ensemble des lois à queue épaisse.

Ci-dessous le tracé des effectifs par tranche de montants qui représente une bonne approximation de la fonction de densité recherchée.



Le calcul du coefficient du skewness donne 7.6 ce qui laisse à penser que la distribution est fortement asymétrique vers la droite.

B. Lois usuelles

Nous allons donc tester quelques lois usuelles à queue épaisse.

1. Loi exponentielle

Dans ce cas, la log vraisemblance s'écrit :

$$\sum \log f(x_i, \alpha, \beta) = n * \log(\alpha) - \alpha * \sum \log(x_i)$$

Les méthodes d'optimisation numériques donnent $\hat{\alpha}=0.000159$ mais le test est rejeté avec une statistique de test de KS de 0.37.

2. Loi gamma

L'expression de la log vraisemblance est

$$\sum \log f(x_i, \alpha, \beta) = \beta * \sum x_i + (\log(\tau(\alpha)) + \alpha * \log(\beta)) + (\alpha - 1) * \sum \log(x_i)$$

Les valeurs obtenues sont $\hat{\alpha}=0.3253$ et $\hat{\beta}=5.1722$ mais le test est rejeté avec une statistique de test de KS de $0.11 > 0.002$ (5% valeur du KS).

Les méthodes d'optimisation numériques donnent $\hat{\alpha}=0.3664093$ et $\hat{\beta}=0.0000532$ mais le test d'adéquation rejette le modèle.

3. Loi log normale

La log vraisemblance s'exprime de la façon suivante :

$$\sum \log f(x_i, \sigma, \mu) = \sum \log(x_i) + n * (\log(2 * \sqrt{\pi}) + \log(\sigma)) + \sum (\log(x_i) - \mu)^2 / (2 * \sigma^2)$$

Les valeurs obtenues sont $\sigma = 1.87$ et $\hat{\beta} = 7$ mais le test est rejeté avec une statistique du test de KS de $0.04 > 0.002$ (5% value du KS).

4. Loi de Burr de type XII

La densité de la loi de Burr à trois paramètres a b s est

$$f(x) = (a b (x/s)^b) / (x [1 + (x/s)^b]^{(a+1)})$$

La maximisation de la log vraisemblance donne les résultants suivants par des méthodes numériques sur R :

a=1.4 b=0.79 s=0.00048 mais là encore les tests ne confirment pas.

5. Loi de Pareto

a. Avec le minimum observé

Le minimum constaté de la distribution est de 30. Nous allons donc tenter de représenter nos montants par une loi de Pareto dont le paramètre d'échelle correspond à cette valeur. L'estimateur du maximum de vraisemblance est dans ce cas complémentèment spécifié. En effet,

Soit $X=(X_1, X_2, \dots, X_n)$ un échantillon aléatoire de loi de Pareto :

$$f(x; \theta) = \theta x^{-(1+\theta)}, \quad 1 \leq x < \infty, \quad \theta \in \mathbf{R}^+$$

Cherchons une estimation par la méthode du maximum de vraisemblance du paramètre inconnu θ

.

Par définition, la fonction de vraisemblance s'écrit :

$$L(\theta; \mathbf{x}) = \prod_{i=1}^n f(x_i; \theta) \equiv D(\mathbf{x}; \theta)$$

Développons cette expression,

$$L(\theta; \mathbf{x}) = \theta^n \prod_{i=1}^n x_i^{-(1+\theta)}$$

En prenant le log,

$$\log L(\theta; \mathbf{x}) = n \log \theta - (1 + \theta) \sum_{i=1}^n \log x_i$$

Dérivons cette dernière expression par rapport à la variable d'intérêt :

$$\begin{aligned} \frac{\partial \log L(\theta; x)}{\partial \theta} &= \frac{n}{\theta} - \sum_{i=1}^n \log x_i = 0 \\ \Rightarrow \hat{\theta} &= \frac{n}{\sum_{i=1}^n \log x_i} \end{aligned}$$

On obtient numériquement $\hat{\theta}=0.2768$

Mais les tests de Kolmogorov Smirnov et de chi deux ne confirment pas avec le calcul d'une p value supérieure à 5%.

b. Avec les méthodes d'optimisation

Nous allons chercher des estimateurs pour les 2 paramètres (échelle et forme). Le log vraisemblance s'exprime de la façon suivante :

$$\sum \log f(x_i, \alpha, \beta) = n * (\log(\alpha) + \alpha * \log(\beta)) - (\alpha + 1) * \sum \log(x_i)$$

Dans ce cas, les estimateurs de la log vraisemblance n'ont pas d'expression explicite. On utilisera alors les méthodes d'optimisation numériques sur le logiciel R.

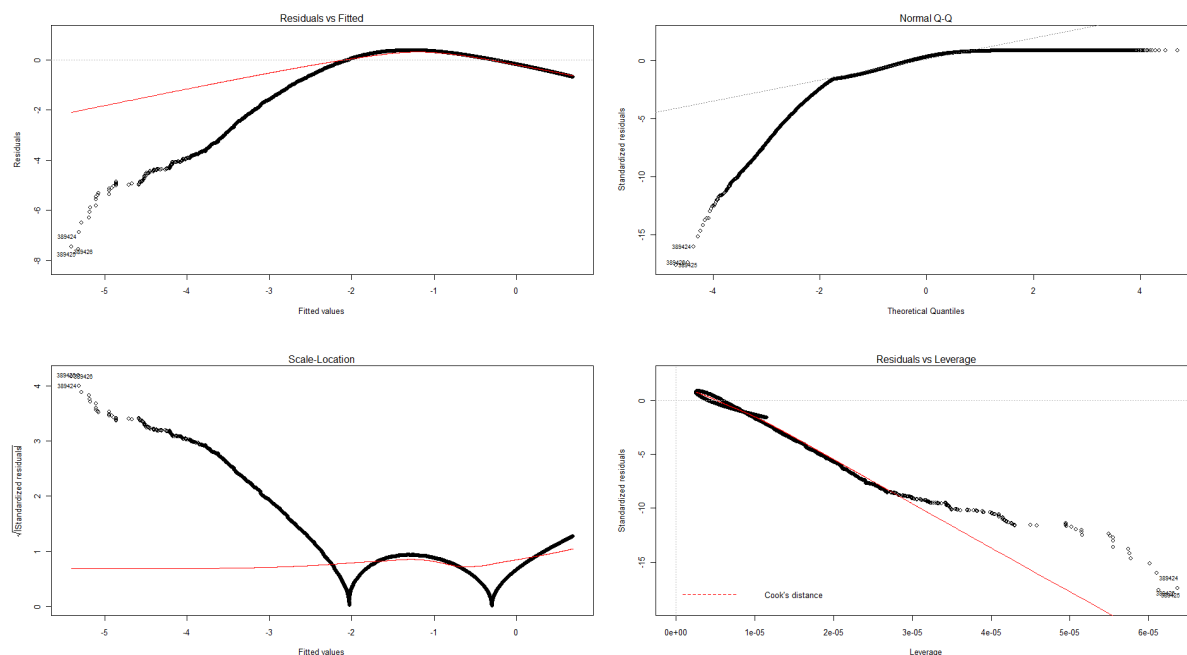
Les valeurs obtenues sont $\hat{\alpha}=0.2975739$ et $\hat{\beta}=0.0000532$ la p value du test de Kolmogorov Smirnov est de 21%.

c. Régression linéaire de la fonction de survie

Dans un modèle parétien, le logarithme de la fonction de survie est une fonction linéaire de la variable. Les premières tentatives ayant échoué, recherchons si il est possible par régression linéaire d'obtenir un modèle Pareto.

La réalisation de cette régression par la méthode des moindres carrés ordinaires donne $\ln(S(x)) = 2.2803 - 0.4676 * \ln(x)$

Analyse des résidus



On constate que la queue de distribution est insuffisamment représentée par la régression. Le comportement de la distribution semble différent pour les grands montants.

C. Le modèle mélange Log normal Pareto

Nous nous sommes intéressés alors aux lois mélanges offrant la possibilité d'alourdir la queue de distribution.

F Planchet et P Théron ont montré dans leur article « Risque de modèle et détermination du capital économique dans le projet Solvabilité 2 » que les indicateurs de queue de distribution (et donc le capital de solvabilité associé) sont perturbés et peuvent être sous estimés lorsque la queue de distribution est insuffisamment prise en compte par la modélisation.

Ils proposent de comparer un modèle log normal validé statistiquement sur des données empiriques avec un modèle équivalent sur presque toute la distribution sauf pour les valeurs extrêmes représentés par une loi de Pareto.

Ils démontrent alors que le capital de solvabilité initial peut être sous estimé malgré la validation de l'adéquation du modèle de base sans queue alourdie.

1. Fonction de survie

La loi est complètement spécifiée par sa fonction de survie :

$$S_X(x) = S_0(x) * 1_{(x \leq m)} + \left(\frac{x}{m}\right)^{-\alpha} * S_0(m) * 1_{(x > m)}$$

Le seuil m de bascule sur la loi de Pareto doit correspondre à un quantile élevé de la distribution de base S_0 .

La loi du phénomène X que l'on souhaite modéliser est une loi de Pareto conditionnellement au fait que le seuil m est dépassé.

Pratiquement, la fonction quantile de X pour des valeurs p inférieurs à $1 - S_0(m)$ est donnée

par :

$$x_p = m * \left(\frac{1-p}{S_0(m)}\right)^{\frac{1}{\alpha}}$$

On a $x_{1-S_0(m)} = m$.

On fixe m en contrôlant $1 - S_0(m)$ à un niveau assez grand mais inférieur à p , typiquement dans un contexte solvabilité $p=99.5\%$ et on choisira $S_0(m)=1\%$ ou 2% .

2. Estimation des paramètres du modèle

Le modèle de base choisi est log normal parce qu'en autres, c'est la loi usuelle dont on a observé le rejet le moins net.

L'estimation des paramètres du modèle est effectuée par la méthode du maximum de vraisemblance. En effet, la log vraisemblance s'écrit en notant $(x_{(1)}, \dots, x_{(n)})$ la statistique d'ordre associée à l'échantillon (x_1, \dots, x_n) et k le plus petit indice tel que $(x_{(k)} \geq m)$

$$l(x_1, \dots, x_n; \mu, \sigma, \alpha) = \sum \ln \left(\frac{1}{\alpha x_{(i)} \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\ln(x_{(i)} - \mu)}{\sigma} \right)^2 \right] \right) + \sum_{i=k}^n \ln(\alpha m^\alpha S_0(m)) x_{(i)}^{-\alpha-1}$$

Ce qui conduit après simplification à :

$$l(x_1, \dots, x_n; \mu, \sigma, m, \alpha) = cste - (k-1) \ln(\sigma) - \frac{1}{2} \sum_{i=1}^{k-1} \left(\frac{\ln(x_{(i)} - \mu)}{\sigma} \right)^2 + (n-k+1) \ln(\alpha) + \alpha(n-k+1) \ln(m) - \alpha \sum_{i=k}^n \ln(x_{(i)}) + (n-k+1) S_0(m)$$

Du fait de la présence de $k = \min \{i; x_{(i)} \geq m\}$ l'expression de la log-vraisemblance est difficilement utilisable en l'état. On peut toutefois décomposer le problème de maximisation en remarquant que

$$\max_{(\mu, \sigma, m, \alpha)} l(x_1, \dots, x_n; \mu, \sigma, m, \alpha) = \max_m \max_{(\mu, \sigma, \alpha)} l(x_1, \dots, x_n; \mu, \sigma, m, \alpha)$$

On note alors qu'une fois m fixé, les expressions des dérivées partielles de la log vraisemblance sont les expressions classiques des deux lois sous-jacentes, sur les plages de données les concernant. Les estimateurs de μ et σ sont ainsi les classiques estimateurs empiriques pour l'échantillon gaussien $(\ln x_{(i)}; i = 1, \dots, k-1)$:

$$\mu = \frac{1}{k-1} \sum_{i=1}^{k-1} \ln x_{(i)} \quad \text{et} \quad \sigma = \sqrt{\frac{1}{k-1} \sum_{i=1}^{k-1} (\ln x_{(i)} - \mu)^2}$$

L'estimateur du paramètre de queue α est fourni par l'expression :

$$\alpha = \frac{n - k + 1}{\sum_{i=k}^n \ln \left(\frac{x_{(i)}}{m} \right)}$$

Il reste à éliminer m , inconnu, dans l'équation ci-dessus. En pratique on pourra donc procéder de la manière suivante :

- On fixe k
- On calcule μ et σ par les estimateurs classiques
- On calcule $m = \exp \left(\mu + \sigma \phi^{-1} \left(\frac{k}{n} \right) \right)$
- L'estimateur (pseudo maximum de vraisemblance) du paramètre de queue α est alors fourni par l'expression $\alpha = \frac{n - k + 1}{\sum_{i=k}^n \ln \left(\frac{x_{(i)}}{m} \right)}$

On obtient une valeur $l(k)$ de la log vraisemblance ; on recommence avec $k' > k$ et on conserve l'estimation des paramètres associée à la valeur maximale de la séquence $l(k)$ ainsi obtenue.

3. Simulation de la loi mélangée

La simulation d'un échantillon issu de la loi mélangée peut être obtenue simplement de la manière suivante :

- Tirage d'une valeur u distribuée uniformément sur $[0,1]$
- Si $u > p_0$, tirage de x dans la loi de Pareto de paramètres (m, α)
- Si $u < p_0$, tirage de x dans la loi $\frac{S_0(x) - S_0(m)}{1 - S_0(m)}$

Ce dernier tirage peut être effectué avec une méthode de rejet : on effectue un tirage dans la loi log normale, et on le rejette si la valeur obtenue est supérieur à m . En effet, comme :

$$P(X > x / X \leq m) = \frac{S_0(x) - S_0(m)}{1 - S_0(m)}$$

Cela conduit exactement à la loi cherchée.

4. Application numérique

Sur nos données de montant on part du 80^e percentile jusqu' au 100^e par pas de 0.1% (soit 200 pas) on sépare à chaque fois la distribution en deux vecteurs. La taille du premier correspond à l'observation du quantile étudié et le second est le complémentaire sur les données.

On calcule les estimateurs classiques de la moyenne et de la variance d'une loi log normale sur ce premier vecteur.

On retient pour seuil de bascule en loi de Pareto, l'inverse de la fonction de répartition de la loi log normale en ce premier vecteur. C'est le paramètre d'échelle de la loi de Pareto, on estime alors le paramètre de forme par maximum de vraisemblance.

On calcule alors la vraisemblance de l'échantillon vis-à-vis de la loi mélange alors totalement spécifiée.

Le pas, le seuil et les estimateurs associés qui maximisent la vraisemblance sont retenus pour qualifier la loi mélange. Le pas qui maximise la vraisemblance est le 199^e pour un quantile de 99.8%

La loi choisie après calcul est :

$$\mu = 7$$

$$\sigma = 1.916$$

$$m = 273\ 030\ \text{€}$$

$$\alpha = 2.4$$

Les montants au-delà de 273 030 suivent une loi de Pareto de paramètre 2.4

5. Test d'adéquation

La statistique de test obtenu par Kolmogorov est $D_n = 0.00177440643408505$

Elle est inférieure (au seuil de 5%) à 0.00217934725835054 .

Le modèle est donc statistiquement validé.

L'intérêt majeur d'avoir pu établir une distribution des résultats de couverture annuels est la possibilité d'étudier les scénarios les plus défavorables en vue de quantifier les risques encourus. Ceci se fait par l'étude des queues de distribution et des indicateurs associés que nous abordons dans la partie suivante.

D. Nouvelle simulation de la distribution des résultats de couverture

Nous pensons au vu des résultats que les valeurs extrêmes sont insuffisamment représentées dans notre modèle en 2008 puisque la queue de distribution n'est que de -800 K€ sur 2008 alors que l'on a constaté empiriquement des amplitudes plus importantes.

Ceci vient du fait que la faible fréquence des grands retards est prise en compte deux fois : naturellement par la loi mélange et une seconde fois par les probabilités de saisie sur une journée, il nous faut donc annuler ce deuxième effet.

Avec la loi des montants décrite plus haut nous générons une matrice de 390 000 lignes de montants et 250 colonnes correspondant à 250 ans (une colonne correspond à un jeu de montants).

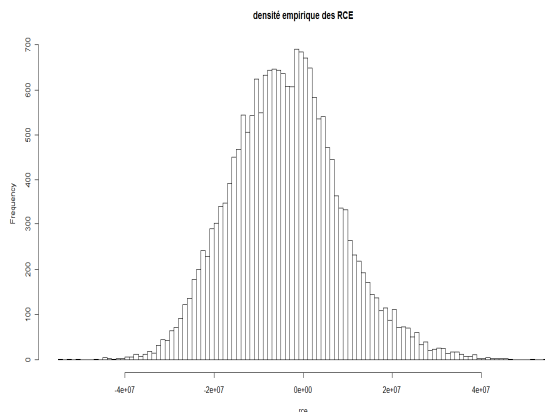
Nous avons retenu le montant du seuil de notre loi de Pareto 273K€ à partir duquel nous agrégerons les flux en deçà en 250 montants dont la somme correspond à la somme de nos flux observés. Pour cela nous réalisons une boucle d'agrégation toutes les 250 lignes. Au delà du seuil, les flux sont conservés tels quels (entre 300 et 400 en moyenne)

La distribution obtenue se caractérise par les quantiles suivants :

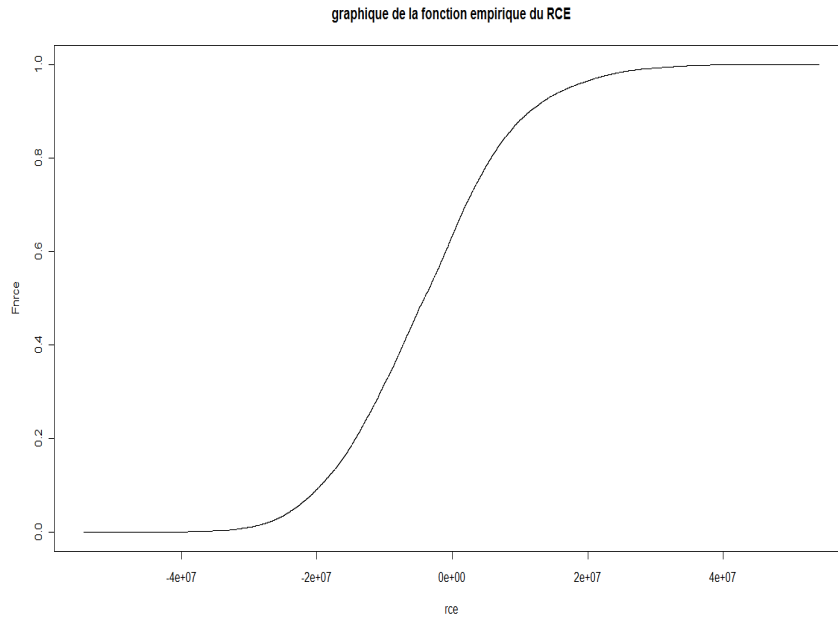
| Nb de scénarios | Min | Moyenne | Quartile 1 | Médiane | Quartile 3 | Max |
|-----------------|----------------|---------------|---------------|--------------|-------------|--------------|
| 10 000 | - 50 066 573 € | - 3 696 748 € | -12 026 606 € | -3 935 957 € | 3 962 067 € | 54 155 756 € |
| 15 000 | - 54 407 032 € | - 3 849 011 € | -12 181 048 € | -4 086 132 € | 3 791 235 € | 54 155 756 € |
| 20 000 | - 54 407 032 € | - 3 908 374 € | -12 348 609 € | -4 172 584 € | 3 746 044 € | 54 155 756 € |

| Nb de scénarios | Ecart type |
|-----------------|--------------|
| 10 000 | 12 315 612 € |
| 15 000 | 12 260 239 € |
| 20 000 | 12 290 705 € |

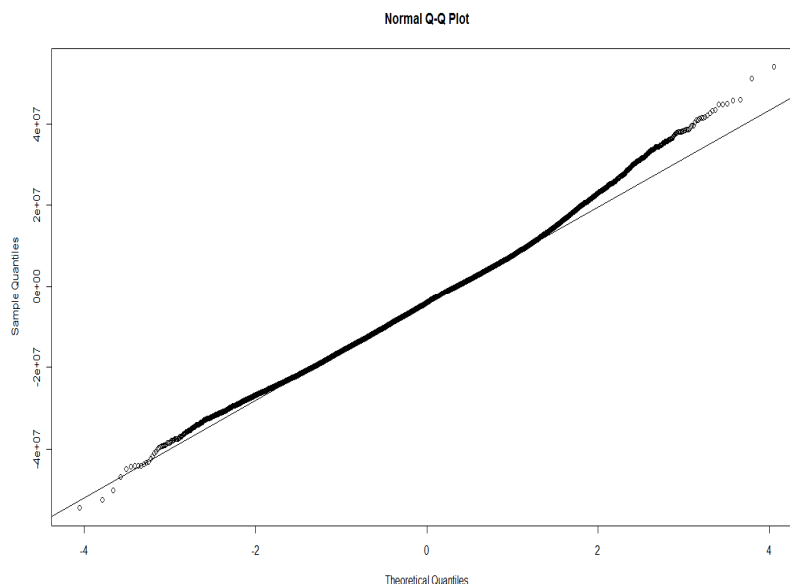
La densité empirique est la suivante :



La fonction de répartition est :



Le calcul du coefficient d'asymétrie (Skewness) qui est proche de 0 soit 0.26 (positif donc légèrement allongé vers droite c'est-à-dire vers les RCE négatifs). En revanche, le coefficient d'aplatissement qui mesure le degré de concentration des observations dans les queues (Kurtosis normalisé) est de $0.34 > 0$. Les queues de distributions ont plus de poids que pour une loi normale. C'est une distribution à queue épaisse. Ceci nous est confirmé par le tracé du qqplot par rapport à la loi normale de moyenne et écart type équivalent.



Par ailleurs le test de Kolmogorov Smirnov a rejeté la normalité :

La statistique $D = 0.0179$ et la p-value nécessaire est de $5.156e-06$.

Nous devons donc nous attacher à l'analyse des valeurs extrêmes de la distribution.

VII. RISK MANAGEMENT ET RESULTAT DE COUVERTURE

La maîtrise des risques est un enjeu majeur en entreprise d'assurances. La capacité des assureurs de faire face à leurs engagements est fondamentalement liée aux montants des fonds propres nécessaires aux sociétés d'assurance.

Ces dernières années, les méthodes régissant les calculs de besoin en fonds propres définies dans le cadre de solvabilité I sont en train d'être totalement revues tant dans l'esprit que dans l'approche calculatoire. Les directives du projet « solvency II » ont vocation à s'appliquer définitivement en 2012. Nous pouvons appréhender nos risques de résultat de couverture (liés aux retards de saisie) sous l'angle de ces deux visions.

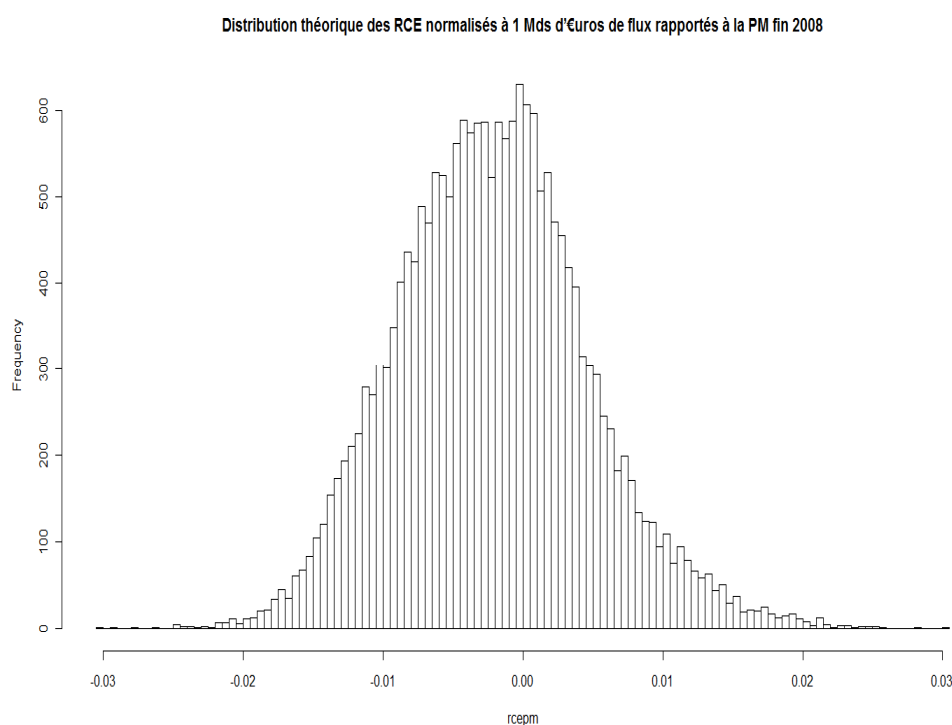
A. Solvabilité I

Le système européen actuellement en vigueur, *Solvency I*, est dit « Fixed Ratio » ou à ratio fixe en fonction des primes, des prestations et des provisions techniques. L'exigence de fonds propres correspondant à la marge réglementaire est essentiellement basée sur 1% des provisions mathématiques en UC.

Les provisions mathématiques de notre portefeuille en UC sont à fin 2008 de **1.8Mds d'€**.

Distribution théorique des RCE 2008 normalisés à 1 Mds d'euros de flux rapportés à la PM fin 2008

Il convient de rapporter les montants de RCE calculés sur le montant de la PM fin 2008.



L'ensemble des RCE se trouve entre -300 points de base et +300 points de base. Certains scénarios n'entrent pas dans le cadre de la marge réglementaire. Les frais de gestion étant globalement de 1%, nous aurions insuffisamment de quoi faire face aujourd'hui. De plus, cette marge de 1% est à mettre au regard de l'ensemble des autres risques de l'assureur. Cette approche en dehors de son aspect couperet, ne nous renseigne pas sur la vision globale des risques supportés par l'assureur, ni sur la corrélation des risques entre eux. Il convient de rappeler à notre lecteur les principes de Solvency II avant de voir leurs applications aux risques étudiés ici.

B. Solvabilité II (Solvency II)

Par rapport à la directive « Solvency I », cette deuxième version propose de mesurer plus finement le risque réel de la compagnie, en laissant le choix du niveau d'adaptation interne (depuis l'approche standard du risque jusqu'au développement de modèles internes). Solvency II généralise la mesure du risque opérationnel, introduit le Solvency Capital Requirement (SCR) et entraîne un contrôle accru du régulateur. La réforme doit permettre d'ajuster les besoins en fonds propres selon la nature des risques : souscription, marché, gestion actif-passif, opérationnels.

Le premier volet de Solvency II accentue l'exigence en termes de quantité et de qualité du capital et instaure deux paliers dans le capital de solvabilité, le SCR et le MRC.

- Le Capital de Solvabilité Requis (SCR) a pour rôle d'absorber les pertes imprévues, en cas d'exercice à forte sinistralité par exemple. Le SCR doit ainsi tenir compte de tous les risques significatifs et quantifiables. En cas de son non respect, la compagnie doit décider d'un plan d'action qui devra être approuvé par le superviseur. Le SCR est calculé via une formule standard ou via un modèle interne élaboré par la compagnie, et doit correspondre à une *Value at Risk* de 99,5%, en d'autres termes à une probabilité de survenance de faillite d'une fois tous les 200 ans.
- Le Minimum de Capital Requis (MCR) est inclus dans le SCR. Il s'agit du montant minimum de fonds propres à détenir sous peine d'intervention du superviseur avec un possible retrait d'agrément. Son calcul devant être « robuste, simple, et auditable », il n'impliquera pas de modèle interne et devrait vraisemblablement correspondre à une *Value at Risk* entre 80 et 90% à un an.

1. Calcul du SCR et value at risk sur la distribution 2008 standardisée à un milliard de flux

On rappelle que par rapport aux orientations retenues dans le QIS4 réalisée par le CEIOPS courant 2008, il est prévu que chaque risque entraîne un besoin en capital, estimé à partir de formules spécifiques destinées à refléter au mieux la nature du risque mesuré. La technique retenue ici est un « scénario based ».

On rappelle que Value at Risk=VaR (99.5%)= $P(RCE \leq Var(99.5\%)) = 0.05$

Le montant de value at risk à 99.5% pour notre distribution de RCE 2008 s'élève à :

| | |
|----------------------------|----------------|
| Value at risk 99,5% | |
| valeur nominale | - 32 719 256 € |
| en % des PM | - 3,272% |

Nous avons donc ici une estimation de la VaR à horizon annuel.

Pour un milliard de flux au global (entrées et sorties), nous devrions donc immobiliser 32 M€ soit 3.3%.

2. Risk management et CTE

La VaR est une mesure du quantile d'une distribution de la variable économique à un niveau de risque donné. De ce fait, elle ignore toutes les informations concernant la queue de la distribution sous-jacente (le risque de queue). La VaR ne mesure donc pas les pertes excédant le seuil de risque fixé. Ainsi, si la distribution empirique présente un potentiel élevé de grandes pertes au-delà du niveau de confiance, la VaR pourrait être une mesure de risque inappropriée. Autrement dit, elle pourrait sous-estimer de manière significative le risque. Pour remédier à ce problème, il existe des mesures de perte moyenne de la queue de distribution.

L'une d'entre elles est le CTE (Conditional Tail Expectation).

L'estimateur CTE a été proposé par Artzner et al² en 1999 comme alternative à la VaR qui est critiquée comme une mesure incohérente. L'avantage du CTE par rapport à la VaR consiste en sa propriété de sous-additivité et son impact en termes d'agrégation de risques. L'estimateur CTE fournit une mesure robuste, pratique et cohérente d'un risque financier au point qu'il est devenu actuellement la mesure préférée d'évaluation du capital financier. Comme son nom l'indique, le CTE mesure l'espérance conditionnelle d'une variable aléatoire au niveau de sa queue de distribution. On a plus précisément la définition suivante:

Définition : Soient donnée X une variable aléatoire réelle et α un seuil de confiance ($\alpha \in]0,1[$, plutôt proche de 0). On a :

$$CTE(\alpha) := E(X | X < q_\alpha),$$

où q_α est le quantile d'ordre α , c'est-à-dire le réel défini par :

$$P(X \leq q_\alpha) = \alpha.$$

Lorsqu'une variable aléatoire est de loi inconnue mais qu'on dispose d'un échantillon aléatoire de taille n de valeurs de X , alors en ordonnant l'échantillon en question par ordre croissant: $x_1 \leq x_2 \leq \dots \leq x_k \leq \dots \leq x_n$, on peut estimer empiriquement le CTE de X au niveau $\alpha = k/n$ par :

² P. Artzner, F. Delbaen, J.M. Eber & D. Heath, Coherent measures of risk, *Mathematical Finance*, 9(3), p. 203-228, 1999.

$$\hat{CTE}(\alpha) = \frac{1}{k} \sum_{i=1}^k x_i .$$

Néanmoins, quand la taille de l'échantillon dont nous disposons n'est pas suffisamment grande ou quand le niveau α est très petit, l'estimation empirique précédente pourrait être assez éloignée du $CTE(\alpha)$ réel. Une façon de remédier à ce problème est de considérer plusieurs échantillons de même taille afin de calculer plusieurs $\hat{CTE}(\alpha)$ et d'obtenir après une estimation de la variance de $\hat{CTE}(\alpha)$. En revanche, la simulation de plusieurs échantillons est coûteuse en termes de temps.

Dans un récent article, B. J. Manistre & G. H. Hancock³ ont donné une formule approchée pour $Var(\hat{CTE}(\alpha))$ qui n'utilise qu'un seul échantillon de la variable aléatoire X considérée. Cette formule est :

$$Var(\hat{CTE}(\alpha)) \approx \frac{Var(x_1, \dots, x_k) + (1 - \alpha)(\hat{CTE}(\alpha) - x_k)^2}{k} .$$

Dans le même article, on en trouve à la fois une justification mathématique et une confirmation numérique sur certains modèles (log-normal, Pareto, etc.).

Cet indicateur calculé sur notre distribution de RCE au même niveau que la VaR soit 99.5% donne le résultat suivant :

| | | |
|--------------------|---|--------------|
| CTE (99,5%) | - | 37 129 837 € |
|--------------------|---|--------------|

Nous en déduisons que la perte moyenne au delà du VaR calculé ne dérape pas de façon inconsidéré, ceci d'autant plus que le calcul de l'écart type associé à la formule d'approximation vu plus haut nous a donné le chiffre suivant :

| | | |
|----------------------------------|--|-----------|
| Ecart-type du CTE (99,5%) | | 609 485 € |
|----------------------------------|--|-----------|

Le CTE n'est donc que faiblement dispersé. Pour affecter un montant moindre de capital, le cadre de Solvency II autorise l'application de modèles internes, mais assorti ce droit à un suivi régulier et à une démonstration de la pertinence du suivi des risques.

Le CTE (65%) a été utilisé dans le groupe AXA comme benchmark pour les produits à annuités garanties « guaranteed annuities ».

| | | |
|--------------------------------|---|--------------|
| CTE (65%) | - | 16 661 028 € |
| Ecart-type du CTE (65%) | | 104 154 € |

L'utilisation de cet indicateur aboutirait à un SCR grosso modo deux fois plus petit soit égal à 1.67% pour un milliard de cash flows en valeur absolue.

³ B. John Manistre & Geoffrey H. Hancock, Variance of the CTE Estimator, *North American Journal*, volume 9, Number 2, pages 129-156, 2005.

Nous avons modélisé le résultat de couverture lié à notre activité. Nous avons calculé les niveaux de SCR attribuables à ces risques selon l'approche Solvency II. Nous avons calculé le CTE au-delà de 65% et estimé son écart type ce qui nous a mis en évidence le fait que la perte restait confinée autour de notre indicateur de solvabilité. Dans la philosophie de Solvency II, les modèles de risque construits ne doivent pas uniquement servir à évaluer une exposition aux risques de façon passive, mais ils doivent aussi être le socle d'une réelle politique de risk management.

Le législateur doit en tirer la garantie d'une maîtrise des risques dynamique et évolutive, tandis que l'assureur peut bénéficier d'une réduction des fonds propres immobilisés. Il doit pour cela mettre en œuvre des actions visant à réduire son exposition aux risques et en démontrer la pertinence par l'intermédiaire de ses modèles.

Nous avons modélisé une loi de retard via un mélange d'exponentielles ou chacune d'elles était liée à une typologie de risque. Par ailleurs, ces trois risques ne sont pas de même nature : la première exponentielle concernait les retards conscients, ils ne résultent pas d'une erreur humaine ou liée aux systèmes d'information et sont assimilables selon la nomenclature Solvency II à des risques de **souscription vie**. En effet, ce dernier englobe l'ensemble des risques lié à une tarification insuffisamment prudente lors de la souscription ou le rachat du contrat. Les coûts liés à cette première loi devraient être pris en compte au niveau de la tarification, ce qui confirme notre choix de segmentation. La troisième exponentielle décrit des erreurs de fonctionnement humain ou technique et cadre donc bien avec la conception du **risque opérationnel** selon Solvency II. Enfin, nous avons constaté les deux types de risque associés à la deuxième exponentielle avec en plus une composante intentionnelle sur certains cas (gestes commerciaux consentis à titre exceptionnel).

Nous allons étudier successivement et individuellement chacun de ces risques et réévaluer nos indicateurs. Au vu de ces derniers, nous discuterons des mesures à prendre ou recommandations à faire au service client.

3. Réévaluation du modèle avec la loi des retards de longue durée

La loi des retards de longue durée est une exponentielle de paramètre $\lambda_1 = 0.014577$.

Nous avons donc de nouveau généré 20 000 scénarios de retards correspondant à cette loi et réévalué la distribution de nos RCEs en réutilisant nos 20000 scénarios financiers, puis nous avons standardisé pour un milliard de flux en valeur absolue.

Nous avons obtenu les résultats suivants :

| Nb de scénarios | Min | Moyenne | Quartile 1 | Médiane | Quartile 3 | Max |
|-----------------|----------------|---------------|---------------|--------------|-------------|--------------|
| 10 000 | - 58 422 213 € | - 4 736 004 € | -15 288 119 € | -5 082 339 € | 5 114 099 € | 62 300 672 € |
| 15 000 | - 68 374 582 € | - 4 930 730 € | -15 626 909 € | -5 219 160 € | 4 882 105 € | 62 300 672 € |
| 20 000 | - 68 374 582 € | - 5 012 056 € | -15 821 388 € | -5 341 266 € | 4 818 836 € | 62 300 672 € |

| Nb de scénarios | Ecart type |
|-----------------|--------------|
| 10 000 | 15 460 785 € |
| 15 000 | 15 433 465 € |
| 20 000 | 15 480 812 € |

Les indicateurs de mesure du risque obtenus sont les suivants :

| VaR (99,5 %) | CTE (65 %) | Ecart-type du CTE (65 %) |
|----------------|----------------|--------------------------|
| - 40 923 744 € | - 21 136 832 € | 129 312 € |

Nous avons vu que les erreurs associées à cette loi s'apparente à **du risque opérationnel**. Au regard des flux 2008, **la mise en place de procédures de contrôle des erreurs humaines et des dysfonctionnements informatiques** doit nous permettre à l'avenir d'économiser du capital non négligeable (évalué à 41 M€ pour un milliard de flux) au titre de cette problématique.

Nous remarquons qu'une fois de plus l'adoption du CTE à hauteur de 65 % allège notre exigence en capital et ceci d'autant plus que l'écart type estimé est relativement faible.

4. Réévaluation du modèle avec la loi des retards de courte durée

La loi des retards de courte durée est une exponentielle de paramètre $\lambda_2 = 1.124861$.

Nous avons également de nouveau généré 20 000 scénarios de retards correspondant à cette loi et réévalué la distribution de nos RCEs en réutilisant 20000 scénarios financiers.

Nous avons obtenu les résultats suivants sur un milliard de flux standardisé en valeur absolue :

| Nb de scénarios | Min | Moyenne | Quartile 1 | Médiane | Quartile 3 | Max |
|-----------------|---------------|-------------|-------------|-------------|------------|-------------|
| 10 000 | - 4 557 182 € | - 147 915 € | - 579 661 € | - 192 234 € | 225 581 € | 4 263 456 € |
| 15 000 | - 4 557 182 € | - 152 272 € | - 580 582 € | - 196 448 € | 219 871 € | 4 263 456 € |
| 20 000 | - 4 557 182 € | - 153 370 € | - 583 543 € | - 197 551 € | 219 630 € | 4 263 456 € |

| Nb de scénarios | Ecart type |
|-----------------|------------|
| 10 000 | 672 481 € |
| 15 000 | 672 381 € |
| 20 000 | 671 787 € |

Les indicateurs de mesure du risque obtenus sont les suivants :

| | |
|----------------------------|---------------|
| value at risk 99,5% | |
| valeur nominale | - 1 811 952 € |
| en % des PM | - 0,181% |

| | |
|------------------|-------------|
| CTE (65%) | - 814 363 € |
|------------------|-------------|

| | |
|--------------------------------|----------|
| Ecart-type du CTE (65%) | 16 472 € |
|--------------------------------|----------|

Le SCR constaté devrait être agrégé dans la **matrice des risques de souscription vie selon la nomenclature Solvency**.

A hauteur d'AXA France, cela correspond pour 2009 à un SCR sur ces risque de

0.2%* 9.5 Milliards d'€uros =19 M€.

Nous recommandons donc d'augmenter les effectifs des équipes de gestion afin de réduire les stock d'actes concernés par les « petits retards » de saisies que nous pensons pouvoir être absorbés par des gestionnaires supplémentaires. Nous rappelons qu'il s'agit pour la plupart d'actes dont le gestionnaire est conscient mais qu'il ne peut accomplir dans les délais nécessaires à la réduction du risque. Des équipes renforcées conduiraient inmanquablement à revoir les paramètres de cette loi afin de diminuer la fréquence de ces retards.

Dans l'absolu, il faudrait arbitrer entre le gain consécutif au gap de rendement de la variation de capital théorique immobilisé induite par ces recrutements et l'augmentation des frais généraux associée (salaires, recrutement...).

On rappelle que la gestion des allers-retours entre les services clients et les middle office commerciaux est une des principales causes des retards de courte durée.

5. Réévaluation du modèle avec la loi des retards de moyenne durée

La loi des retards de moyenne durée est une exponentielle de paramètre $\lambda_3 = 0.099087$.

Nous avons également de nouveau généré 20 000 scénarios de retards correspondants à cette loi et nous avons réévalué la distribution de nos RCEs en réutilisant nos 20000 scénarios financiers.

Les quantiles de la distribution sont les suivants :

| Nb de scénarios | Min | Moyenne | Quartile 1 | Médiane | Quartile 3 | Max |
|------------------------|----------------|----------------|-------------------|----------------|-------------------|--------------|
| 10 000 | - 26 069 763 € | -2 433 870 € | -8 372 455 € | -2 909 914 € | 2 595 450 € | 37 598 618 € |
| 15 000 | - 31 045 253 € | -2 548 754 € | -8 513 160 € | -3 034 976 € | 2 497 894 € | 37 598 618 € |
| 20 000 | - 31 045 253 € | -2 594 209 € | -8 614 806 € | -3 105 470 € | 2 447 240 € | 37 598 618 € |

| Nb de scénarios | Ecart type |
|------------------------|-------------------|
| 10 000 | 8 380 327 € |
| 15 000 | 8 376 929 € |
| 20 000 | 8 411 153 € |

Les indicateurs de mesure du risque obtenus sont les suivants :

| | |
|----------------------------|----------------|
| value at risk 99,5% | |
| valeur nominale | - 20 072 209 € |
| en % des PM | - 2% |

| | |
|------------------|----------------|
| CTE (65%) | - 11 139 626 € |
|------------------|----------------|

| | |
|--------------------------------|----------|
| Ecart-type du CTE (65%) | 61 847 € |
|--------------------------------|----------|

Nous constatons des indicateurs de queue de distribution (pour les cash-flows globaux) dont l'ordre de grandeur est équivalent à celui de la loi des retards de grande durée ou de petite durée. Ce type de retard contenant les deux natures de risque, les mêmes solutions proposées plus haut tendrait à réduire notre exposition (procédures de contrôle plus poussées et recrutement approprié.)

Conclusion

Nous sommes arrivés au terme de notre étude sur les risques de résultat de couverture liés aux actes de gestion des services clients pour un important portefeuille d'assurance vie.

Nous avons cartographié ces risques et nous les avons modélisé par une loi mélange de trois exponentielles conforme à la typologie des risques constatée empiriquement.

Notre approche prudentielle nous a amené à représenter l'évolution de nos supports financiers par un modèle RSLN2 capable de tenir compte des scénarios de crise pour les actifs les plus risqués (marchés actions).

Nous avons ainsi établi la loi de nos résultats de couverture annuels et démontré la convergence vers cette loi lors de l'utilisation de plusieurs scénarios de retards de saisie et d'évolution financière. Nous avons démontré la stabilité de cette distribution pour un Milliard d'€uros de flux totaux lorsque les rapports relatifs des cash-flows positifs et négatifs sont stables. Ceci nous a permis de nous intéresser à la quantification des risques encourus par l'étude de la queue de distribution.

Nous avons utilisé la norme Solvency II, soit l'indicateur Value at risk à 99.5%.

Nous avons démontré que l'exigence en capital est de 37 M€.

Par ailleurs, nous avons introduit l'indicateur CTE dont nous soulignons les propriétés par rapport à la VaR, nous retenons le niveau de 65% de benchmark utilisé chez AXA pour la tarification de produits à annuités garanties.

Dans ces conditions le SCR peut être ramené à 16 M€.

Les SCR calculés ici peuvent être associés aux autres modèles stochastiques constitutifs du modèle interne en cours d'élaboration chez AXA.

Nous avons ensuite affiné l'analyse de chacune des composantes de notre loi de retard en remarquant que la nature des risques n'est pas la même (opérationnel, souscription vie).

Nous avons pu démontrer la prépondérance du risque opérationnel dû aux erreurs de saisie humaine ou aux dysfonctionnements des systèmes d'information. (VaR (99.5%) de 7 à 8.3%).

Hormis l'évaluation quantitative, nous respectons parfaitement ici la philosophie de Solvency II en mettant en évidence les nécessaires gestions et pilotage de ces risques financiers, ceci d'autant plus qu'aucun compte de résultat ou provision ne leur est affecté. En effet, la « prise en charge de ce risque » est implicitement contenue dans les frais de gestion ou frais relatifs aux actes (entrée, arbitrage) qui ont tendance à être fortement réduits sous la pression concurrentielle.

Par ailleurs, nous avons fait des recommandations afin de réduire notre exposition aux risques (arbitrages couts/produits pour les recrutements aux services clients) et accroissement des contrôles pour le risque opérationnel. La gouvernance des risques peut encore être améliorée en les étudiant de façon encore plus affinée. Il est par exemple clairement établi pour notre périmètre que des saisonnalités existent dans la demande d'actes (chute du nombre d'actes entrants en période estivale et augmentation au mois de décembre afin de bénéficier de l'exercice fiscal), or celle-ci a trivialement un lien direct avec le stock d'actes à terminer à effectifs constants. Il est probablement judicieux de renforcer les équipes service clients en tenant compte de cette saisonnalité.

Au delà de cette remarque, la qualité de service s'en trouvera bien évidemment améliorée. Il conviendrait alors de réajuster les paramètres de notre loi à posteriori, ce qui devrait nous permettre de réduire encore les exigences réglementaires afférentes à ces risques.

Il nous reste cependant à valider notre modélisation auprès d'autres services clients du groupe AXA (non fait à ce jour), l'accès aux données de nos concurrents étant bien évidemment impossible.

Par ailleurs, d'autres types de risques opérationnels ayant un impact direct sur le résultat de couverture n'ont pu être modélisés ici : erreurs ou dysfonctionnements dans le passage d'ordre, problématique des process complexes impactant le résultat de couverture (détachements de coupons rétroactifs, opérations d'arbitrage suite à fusion-absorptions, gestion sous mandat, liquidation...) qui constituent des risques opérationnels nécessitant une approche théorique à part entière.

Bibliographie

Cours de statistique mathématique. A MONFORT 2001

Cours de finance 2^e année CEA JL AURE 2002

Processus stochastiques et fiabilité des systèmes C. COCCOZA-THIVENT 1997

A regime switching model of long term stock returns M HARDY M.R., 2001 North American Actuarial Journal

Variance of the CTE estimator MANISTRE & HANCOCK: North American Actuarial Journal, Vol. 9, No. 2, pp.129-156, 2005

Cours de séries temporelles C GOURIEROUX 1983

Nonparametric statistics for the behavioral sciences, S.SIEGEL, N.J. CASTELLAN 1988

Risque de modèle et détermination du capital économique dans le projet Solvabilité
P. THEROND & F PLANCHET (article en ligne page personnelle de Mr Thérond) 2007

Provisions techniques et capital de solvabilité d'une compagnie d'assurance : méthodologie
d' utilisation de Value-at-Risk P. THEROND & F PLANCHET Revue Assurances et
gestion des risques 2007

Annexes

Programmes SAS

Tests wilcoxon

```
proc nparlway data = distribution wilcoxon;
class annee;
var rce;
exact;
title 'test de Wilcoxon & Kruskal-Wallis ';

*output ;
run ;
proc nparlway data = distribution edf;
class annee;
var rce;
title 'test de KS';

*output ;
run ;
```

Programmes sur R (loi log normale pareto)

```
log_vraisemblance=function(data, j, s){

  k=length(which(data<=quantile(data, s+(1-s)*(j-1)/200)))
  data_min=as.vector(1:(k-1))
  data_max=as.vector(1:(n-k+1))
  for(i in 1:(k-1)){
    data_min[i]=data[i]
  }
  for(i in k:n){
    data_max[i-k+1]=data[i]
  }
  mu=sum(log(data_min))/(k-1)
  sigma=(sum((log(data_min)-mu)^2)/(k-1))^(1/2)
  seuil=exp(mu+sigma*qnorm(k/n, 0, 1))
  alpha=(n-k+1)/sum(log(data_max/seuil))
  logvrai=sum(log(dlnorm(data_min, mu,
sigma)))+sum(log(alpha*(seuil^alpha)*plnorm(seuil, mu,
sigma, lower.tail=FALSE)*(data_max^(-alpha-1))))

  res=as.vector(1:5)
  res[1]=mu
  res[2]=sigma
  res[3]=seuil
  res[4]=alpha
  res[5]=logvrai
```

```

        res
    }
    lnPareto=function(mu,sigma,seuil,alpha,x) {
#Densité de la loi composée log-normale / Pareto

        if (x<seuil) {
            p=dlnorm(x,mu, sigma)
        }
        else {

            p=alpha*seuil^alpha*plnorm(seuil,mu,sigma,lower.tail=FALSE)*x^(-
alpha-1)
        }

        p
    }

}

#Déclaration bibliothèques
source("log_vraisemblance.r")
source("lnPareto.r")
#Lecture des données

#Initialisations
logvrai=as.vector(1:200)
s=.8
n = length(data)

#Calcul de la log-vraisemblance pour différents seuils
for(j in 1:200){
    logvrai[j]=log_vraisemblance(data,j,s)[5]
}

#Identification de la valeur maximum
j_hat=which.max(logvrai)
#Affectation des valeurs de l'estimateur
res=log_vraisemblance(data,j_hat,s)
#Affichage
print(paste("mu=",res[1]))
print(paste("sigma=",res[2]))
print(paste("seuil=",res[3]))
print(paste("alpha=",res[4]))
write(c("mu", "=", res[1]),file="resQ2
pareto.txt",append=TRUE,ncolumns=3)
write(c("sigma", "=", res[2]),file="resQ2
pareto.txt",append=TRUE,ncolumns=3)
write(c("seuil", "=", res[3]),file="resQ2
pareto.txt",append=TRUE,ncolumns=3)
write(c("alpha", "=", res[4]),file="resQ2
pareto.txt",append=TRUE,ncolumns=3)
lognormal
lognormal
lognormal
lognormal

#Définition de la densité de la loi ajustée
ll=function(x) {lnPareto(res[1],res[2],res[3],res[4],x)}

#Construction du graphique
abscisse=as.vector(1:200)
for(i in 1:200){

```

```

    abscisse[i]=(i-1)*(1-s)/200+s
  }

plot(abscisse,logvrai,type="l",xlab='Probability',ylab='log-likelihood')

#Distribution ajustée
x=as.vector(1:1000)
y=as.vector(1:1000)
for(i in 1:1000){
  x[i]=min(data)+i*(max(data)-min(data))/1000
  y[i]=ll(x[i])
}
plot(x,y,type="l",xlab='Value',ylab='Density')

probpareto=function(x,alpha,seuil) {
#fct de repartition de la loi composée Pareto / Pareto

  p=1-((x/seuil)^(-alpha))

  p
}

invpareto=function(p,alpha,seuil) {
#inverse de la fct de repartition loi composée Pareto / Pareto

  xp=exp((alpha*log(seuil)-log(1-p))/alpha)

  xp
}

glnPareto=function(mu,sigma,seuil,alpha,nb) {
#generateur de la loi composée log-normale / Pareto

  x <-numeric(n)
summary(x)

  p=runif(n, min=0, max=1)

for (i in 1:length(data))

{
  if (p[i]>plnorm(seuil,mu,sigma))

  {
    x[i]=invpareto(runif(1,min=0, max=1),alpha, seuil)
  }

else {
  x[i]=qlnorm((1-p[i])*plnorm(seuil,mu,sigma),mu,sigma)}
}
x

```

```

}
}

glnPareto(res[1],res[2],res[3],res[4],nb)

x <-numeric(n)
summary(x)

      p=runif(n, min=0, max=1)
#p=numeric(nb)

for (i in 1:length(data))

{
  if (p[i]>plnorm(res[3],res[1],res[2]))

  {
    x[i]=invpareto(runif(1,min=0, max=1),res[4],res[3])
  }

  else {
    x[i]=qlnorm((runif(1,min=0,
max=1))*plnorm(res[3],res[1],res[2]),res[1],res[2]) }
}

x=sort(x)
summary(x)
#####
##### Calcul et Graph de Fnx #####
#####
Fn=c()
for (i in 1:n){
j=i
while(j<n && data[j]==data[j+1]){
j=j+1
}
Fn[i]=j/n
}

###Calcul
#Les données sont classées
Fnx=c()
for (i in 1:n){
j=i
while(j<n && x[j]==x[j+1]){
j=j+1
}
Fnx[i]=j/n
}
###Graph
graph_Fnx<-function(){

plot(data,Fnx, type = "l", col = "black", lwd = 1, main = "graphique de la
fonction empirique de la loi estimée log normale pareto")
}

```

```

#####
##### Test KS #####
#####

diff1=abs(Fn-Fnx)
diff2=c()
diff2[1]=abs(Fnx[1])
diff2[2:n]=abs(Fn[1:(nb-1)]-Fnx[2:nb])
max1=max(diff1)
max2=max(diff2)
Dn=max(max1,max2)
calpha=1.36/sqrt(n)
write(c("Dn", "=", Dn), file="resQ2          log          normal
pareto.txt", append=TRUE, ncolumns=3)
write(c("c_{alpha}", "=", calpha), file="resQ2  log  normal  pareto  sup
30.txt", append=TRUE, ncolumns=3)
if (Dn > calpha){
write("On rejette le modele log normal pareto", file="resQ2 log normal
pareto .txt", append=TRUE)
}else{
write("On ne rejette pas le modele log normal pareto", file="resQ2 log
normal pareto .txt", append=TRUE)
}

#####
##### generation matrice 250 ans loi #####
#####

mat <- matrix(0, nrow = n, ncol=250)

for (j in 1:250)
{
p=runif(n, min=0, max=1)
for (i in 1:n)

{
if (p[i]>plnorm(res[3],res[1],res[2]))

{
mat[i,j]=invpareto(runif(1,min=0, max=1),res[4],res[3])
}

else {
mat[i,j]=qlnorm((runif(1,min=0,
max=1))*plnorm(res[3],res[1],res[2]),res[1],res[2]) }
}

}

}

```

Programmes VBA fonction log vraisemblance RSLN2

```
Function LogLJ(P_11 As Double, P_22 As Double, Mu_1 As Double, Mu_2 As Double,
Sigma_1 As Double, Sigma_2 As Double) As Double
Dim i As Integer
Dim M1() As Double      'correspond à matdist( , 1)
Dim M2() As Double      'correspond à matdist( , 2)
Dim X1() As Double      'correspond à matx( , 1)
Dim X2() As Double      'correspond à matx( , 2)
Dim L As Double         'valeur de retour de la fonction LogLJ
Dim X1prev As Double
Dim X2prev As Double
Dim f As Double
Dim P_12 As Double, P_21 As Double

P_21 = 1 - P_22
P_12 = 1 - P_11

ReDim M1(UBound(Y))
ReDim M2(UBound(Y))
ReDim X1(UBound(Y))
ReDim X2(UBound(Y))

For i = 0 To UBound(Y)
M1(i) = DensNor(Y(i), Mu_1, Sigma_1)
M2(i) = DensNor(Y(i), Mu_2, Sigma_2)
Next i

X1(0) = Pi_1(P_11, P_22) * M1(0)
X2(0) = Pi_2(P_11, P_22) * M2(0)
L = Log(X1(0) + X2(0))

For i = 1 To UBound(Y)
f = X1(i - 1) + X2(i - 1)
X1prev = X1(i - 1) / f
X2prev = X2(i - 1) / f
X1(i) = M1(i) * (P_11 * X1prev + P_21 * X2prev)
X2(i) = M2(i) * (P_12 * X1prev + P_22 * X2prev)
L = L + Log(X1(i) + X2(i))
Next i

LogLJ = L
End Function
```