

Mémoire présenté le :
pour l'obtention du Diplôme Universitaire d'actuariat de l'ISFA
et l'admission à l'Institut des Actuares

Par : Gaylord LEGRIS

Titre : Analyse de Sensibilités sur des Portefeuilles de Passifs en Assurance Vie selon les indicateurs Solvabilité 2 : Approche par Générations et Agrégations des portefeuilles

Confidentialité : ☒ NON ☐ (Durée : ☐ 1 an ☐ 2 ans)

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

*Membres présents du jury de Signature
l'Institut des Actuares*

Entreprise :

Nom :

Signature :

*Directeur de mémoire en entre-
prise :*

Nom :

Signature :

*Membres présents du jury de
l'ISFA*

Invité :

Nom :

Signature :

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

Signature du responsable entreprise

Signature du candidat

Table des matières

Résumé	iv
Abstract	v
Remerciements	vi
Synthèse	vii
Synthesis	viii
Introduction	1
1 Contexte réglementaire et modélisation en assurance vie	3
1.1 Les spécificités des produits d'assurance vie épargne	4
1.1.1 Principes fondamentaux du contrat d'assurance vie	4
1.1.2 Les principaux supports d'investissement	5
1.2 Le cadre prudentiel Solvabilité II	6
1.2.1 Le Pilier 1 : Exigences quantitatives et Bilan Prudentiel	7
1.2.2 Le Pilier 2 : Exigences qualitatives et gouvernance	10
1.3 Les Générateurs de Scénarios Économiques (GSE)	10
1.3.1 L'univers Risque Neutre (\mathbb{Q}) : un cadre pour la valorisation	11
1.3.2 L'univers Monde Réel (\mathbb{P}) : un outil de pilotage stratégique	13
1.3.3 Synthèse des deux univers	13
1.4 La représentation du passif : le concept de <i>Model Point</i>	14
1.4.1 La nécessité de l'agrégation	15
1.4.2 Les impacts de l'agrégation en MP sur les indicateurs S2	15
2 La Gestion Actif-Passif et le Modèle de Simulation en Python	16
2.1 Principes et Enjeux de la Gestion Actif-Passif (ALM)	17
2.1.1 La Modélisation ALM : un Outil de Projection Essentiel	17
2.2 Architecture et Fonctionnement du Modèle de Projection	18
2.2.1 Phase 1 : Initialisation du Modèle	19
2.2.2 Phase 2 : La Boucle de Projection Annuelle	21
2.2.3 Phase 3 : Finalisation et Génération des Outputs	22

2.3	Limites Actuelles du modèle	23
3	Construction d'un Générateur de Portefeuilles de Passifs Réaliste	24
3.1	Objectifs Stratégiques et Contraintes Techniques	25
3.1.1	Définition du générateur de portefeuilles de passifs	25
3.1.2	Besoins métiers : simulation de nouveaux produits et analyse concurrentielle	25
3.1.3	Défis de la modélisation : réalisme, volumétrie et flexibilité	27
3.2	Méthodologie de Génération et Modélisation Statistique	27
3.2.1	Approche stochastique par lois de probabilité	28
3.2.2	Synthèse et Génération du Portefeuille Final	37
3.2.3	Calibration des distributions marginales à partir des données de marché	38
3.2.4	Perspective : modélisation des dépendances par la théorie des copules	39
3.3	Présentation de l'Outil et du Portefeuille de Référence Généré	40
3.3.1	Architecture de l'application en Python et choix technologiques	41
3.3.2	Description des paramètres d'entrée et des formats de sortie	41
3.3.3	Analyse descriptive du portefeuille de référence	41
4	Protocole d'Analyse : Sélection d'une Méthode d'Agrégation	42
4.1	Présentation des Méthodes d'Agrégation candidates	43
4.1.1	Approches par clustering (K-means, DBSCAN/HDBSCAN)	44
4.1.2	Autres approches (MP par âge, MP Mémoire Amine Ben Fadhel, etc.)	44
4.2	Définition du Protocole de Test Comparatif	44
4.2.1	Constitution des portefeuilles de test	45
4.2.2	Définition des critères de sélection : fidélité des indicateurs (BE/SCR), performance et temps de calcul	45
4.3	Analyse Comparative et Choix de la Méthode Optimale	45
4.3.1	Synthèse des performances pour chaque méthode candidate	46
4.3.2	Justification du choix de la méthode retenue pour l'analyse de sensibilité	46
5	Analyse de Sensibilité des Indicateurs S2 aux Caractéristiques du Passif	47
5.1	Définition des Scénarios de Sensibilité	48
5.1.1	Création des portefeuilles de test via le générateur	49
5.1.2	Description des chocs sur les variables clés (âge, montant de la PM, etc.)	49
5.1.3	Scénario d'intégration d'un nouveau produit dans le portefeuille	49
5.2	Analyse de l'Impact de l'Agrégation sur la Mesure des Chocs	49
5.2.1	Comparaison des indicateurs S2 sur portefeuilles choqués granulaires et agrégés	50
5.2.2	Analyse de la fidélité de la méthode d'agrégation à retranscrire la sensibilité	50
5.3	Interprétation des Résultats et Validation de l'Approche	50
5.3.1	Validation de la performance de la chaîne de modélisation (Générateur - Agrégation - Modèle ALM)	51
5.3.2	Enseignements sur la sensibilité des portefeuilles aux modifications de caractéris- tiques du passif	51
6	Conclusion	52

6.1	Résumé des résultats	53
6.1.1	Synthèse des principaux résultats obtenus	53
6.1.2	Impact des méthodes d'agrégation sur les portefeuilles de passifs	53
6.2	Perspectives d'amélioration	53
6.2.1	Axes d'amélioration pour les générateurs de portefeuilles de passifs	54
6.2.2	Évolutions possibles des méthodes d'agrégation et de modélisation ALM	54
6.2.3	Autres domaines d'application des générateurs de portefeuilles de passifs	54
6.3	Conclusion Générale et Perspectives	54
6.3.1	Perspectives et Finalisation des Travaux de Recherche	55
6.3.2	Bilan de l'expérience professionnelle chez Accenture : un levier pour le mémoire . .	56
Annexes		58
Bibliographie		59

Résumé

Petit résumé en français de mon mémoire ?

Abstract

It's a brief sum up in english !

Remerciements

Pour la réalisation de ce mémoire, je tiens à adresser mes remerciements à toutes les personnes qui ont contribué de près ou de loin à son élaboration. Que ce soit au travail avec Lionel Aldeberd qui m'a appris énormément, Lucas Blancheton qui a toujours été là pour m'aiguiller dans mes choix. Merci à Luc pour les relectures et à tous mes collègues stagiaires et alternants, Antoine, Cindy, Eleonore, Lucile, Malak, Manon, Nicolas, Tanguy et Titouan pour leur bonne humeur ! J'ai vraiment passé des très chouettes moments à vos côtés !

Synthèse

Un long résumé en français de mon mémoire

Synthesis

Un long résumé en anglais de mon mémoire

Introduction

Placement privilégié des épargnants français, l'assurance vie a atteint un encours record de 1 989 milliards d'euros à fin 2024 (France Assureurs) [6], confirmant son rôle prépondérant dans le patrimoine financier national. Cette performance est portée par une collecte nette annuelle de +29,4 milliards d'euros, une hausse de 28,2 milliards par rapport à 2023, ce qui témoigne d'une forte attractivité du produit dans un contexte économique incertain. La dynamique de l'année 2024 est marquée par une hausse globale des cotisations de +14 %, bénéficiant tant aux supports en euros (+17 %) qu'aux unités de compte (+8 %). Simultanément, les prestations versées aux assurés sont en recul de 5 %. Cette combinaison d'une collecte dynamique et de prestations maîtrisées place la gestion actif-passif (ALM) au cœur des enjeux stratégiques pour les assureurs, qui doivent piloter cet afflux de capitaux tout en maintenant l'équilibre entre sécurité et rendement pour les épargnants.

Le secteur de l'assurance vie en France est ainsi confronté à un besoin important de pilotage via la gestion actif-passif (ALM). Pour cela, les assureurs s'appuient sur l'utilisation de modèles ALM. Ces modèles simulent l'impact de différentes stratégies ce qui nécessite cependant un grand nombre de projections stochastiques, cela engendre une contrainte opérationnelle majeure : le temps de calcul. Cette contrainte limite non seulement la capacité à explorer en profondeur l'ensemble des risques et des opportunités, mais freine également l'agilité stratégique et la réactivité des prises de décision. C'est de la rencontre entre ces exigences et des contraintes opérationnelles qu'ont les assureurs qu'est née la problématique de ce mémoire : comment concilier la nécessité de rapidité des calculs avec l'impératif de fidélité des indicateurs de risque ? Ce mémoire se propose d'investiguer cette problématique en étudiant l'impact des techniques d'agrégation du passif, notamment par la création de *model points*, c'est-à-dire une manière simplifiée de représenter les contrats d'assurance. L'enjeu est de déterminer si cette modélisation simplifiée des portefeuilles de contrats peut constituer une approximation fiable pour le pilotage stratégique, et d'évaluer sous quelles conditions une telle simplification est valide sans masquer des dynamiques de risque essentielles au pilotage de l'entreprise.

Pour répondre à cette problématique, ce mémoire adoptera une double approche. Premièrement, il s'agira de développer un générateur de portefeuilles de passif puis le reste de l'analyse portera sur les effets de l'agrégation sur un portefeuille représentatif du marché français. L'objectif est de comprendre comment les risques évoluent à travers une agrégation. Ce mémoire ne se contentera pas d'analyser l'impact d'une seule méthode d'agrégation ; au contraire, plusieurs méthodes et approches seront testées. Le critère de sélection de la méthode la plus pertinente reposera sur un triple objectif : minimiser l'écart des indicateurs clés de Solvabilité II (notamment le Best Estimate et le SCR), optimiser la rapidité des calculs

et atteindre le plus haut niveau d'agrégation possible qui garderait une significativité économique pour l'assureur suffisante. L'axe principal de ce mémoire consistera donc à mener une analyse de sensibilités approfondie sur ces portefeuilles, qu'ils soient granulaires ou agrégés. Notre étude s'appuiera sur des indicateurs quantitatifs clés issus de la norme Solvabilité II, en évaluant notamment l'impact des chocs économiques sur le Best Estimate, le Solvency Capital Requirement (SCR) et la Present Value of Future Profits (PVFP). Ces métriques permettront de mesurer rigoureusement comment l'agrégation modifie la perception du risque et la valeur économique du portefeuille.

Ce mémoire s'articulera en un parcours logique et progressif en cinq temps. La première partie posera le cadre conceptuel de l'étude en explorant le contexte réglementaire de Solvabilité II, les produits d'épargne en assurance vie et les fondements des Générateurs de Scénarios Économiques (GSE). Après l'explication du socle théorique, la deuxième partie abordera les principes de la Gestion Actif-Passif (ALM), en détaillant l'architecture du modèle de projection qui servira de base aux travaux de ce mémoire. La troisième partie permettra de poser les bases de l'analyse, avec l'élaboration d'un générateur de portefeuilles de passifs réalistes destiné à produire les données synthétiques cohérentes. Le cœur méthodologique sera présenté en quatrième partie, à travers un protocole d'analyse comparant diverses méthodes d'agrégation afin de choisir celle qui sera utilisée. Enfin, la cinquième partie sera consacrée à l'interprétation des résultats d'agrégations sur des chocs économiques où, par le biais d'analyses de sensibilité approfondies, l'impact de la méthode d'agrégation retenue sur la mesure du risque sera quantifié, validant ainsi la pertinence de l'approche.

Chapitre 1

Contexte réglementaire et modélisation en assurance vie

1.1 Les spécificités des produits d'assurance vie épargne

1.1.1 Principes fondamentaux du contrat d'assurance vie

L'assurance vie est une convention par laquelle un assureur, en contrepartie du versement de primes, s'engage à verser un capital ou une rente à la survenance d'un événement incertain lié à la durée de la vie humaine. Cet événement, qui constitue l'aléa au cœur du contrat, peut être le décès de l'assuré avant une date donnée ou, à l'inverse, sa survie jusqu'à cette date. Ce mécanisme repose sur un cycle de production inversé : l'assureur perçoit les primes bien avant de devoir potentiellement régler les prestations, ce qui l'amène à investir ces sommes sur des horizons de temps longs pour honorer ses engagements futurs.



FIGURE 1.1 – Cycle de production inversé en assurance vie (**graphique temporaire**)

La nature de ces engagements répond à des objectifs variés. Les contrats en cas de vie prévoient le versement d'un capital ou d'une rente à une échéance prévue si l'assuré est en vie ; ils sont typiquement utilisés pour se constituer un complément de retraite ou une épargne de précaution. À l'opposé, les contrats en cas de décès garantissent le versement d'un capital ou d'une rente au(x) bénéficiaire(s) désigné(s) au décès de l'assuré, souvent pour protéger des proches ou anticiper des droits de succession. Il existe également des contrats mixtes qui combinent ces deux garanties.

Le fonctionnement de ces contrats repose sur la capitalisation : les primes versées sont investies pour financer la propre couverture future de l'assuré. De par leur nature, ces engagements s'étendent sur de très longues périodes. Une caractéristique fondamentale de l'assurance vie française est sa liquidité. L'assuré dispose de la possibilité de récupérer son épargne à tout moment via un rachat, qui peut être partiel ou total. Cette faculté de rachat constitue une option dont la valeur et le risque doivent être finement gérés par l'assureur, car son exercice a un impact direct sur les besoins de liquidité du portefeuille. La fiscalité joue un rôle incitatif majeur car les plus-values sont imposées plus lourdement si le rachat intervient

avant la huitième année du contrat. Ceci encourage alors l'épargne de long terme.

La gestion de ces engagements de long terme amène l'assureur à proposer différentes modalités d'investissement. Celles-ci permettent de répartir le risque financier entre l'assuré et l'assureur, définissant ainsi le profil de rendement potentiel du contrat. Un contrat plus sûr aura des possibilités de rendements plus faible qu'un contrat risqué. L'épargne des assurés peut ainsi être investie sur deux principaux types de supports aux profils de risque distincts, qui peuvent être combinés au sein de différents types de contrats.

1.1.2 Les principaux supports d'investissement

L'épargne des assurés peut être investie sur deux principaux types de supports aux profils de risque distincts, qui peuvent être combinés au sein de différents types de contrats.

Le **fonds en euros** est le support historique et sécuritaire de l'assurance vie française. Le risque financier y est intégralement porté par l'assureur, qui s'appuie sur une politique d'investissement prudente, majoritairement orientée vers des actifs obligataires. La sécurité de ce support repose sur un ensemble de garanties contractuelles et réglementaires :

- **La garantie du capital** : C'est la garantie la plus fondamentale. L'assureur garantit à tout moment le capital net investi par l'épargnant. Quelle que soit l'évolution des marchés financiers, la somme initialement versée (nette de frais) ne peut pas diminuer.
- **Le taux technique** : Il s'agit d'un taux de revalorisation minimal garanti sur toute la durée du contrat. Fixé à la souscription, il est aujourd'hui très faible, voire nul, en raison des contraintes réglementaires.
- **Le Taux Minimum Garanti (TMG)** : Plus courant aujourd'hui que le taux technique, le TMG est un taux de rendement minimal que l'assureur s'engage à verser pour l'année à venir. Il est fixé annuellement et permet à l'assureur d'ajuster sa politique de rendement.
- **L'effet cliquet** : Ce mécanisme assure que les intérêts générés chaque année sont définitivement acquis. Une fois distribués, ils s'ajoutent au capital garanti et produisent à leur tour des intérêts les années suivantes. Il est impossible de revenir sur les revalorisations passées.
- **La Participation aux Bénéfices (PB)** : L'assureur a l'obligation légale de redistribuer aux assurés au minimum 85% de ses bénéfices financiers et 90% de ses bénéfices techniques. Cette participation constitue la majeure partie du rendement annuel, au-delà du TMG. Pour lisser les performances, une partie de cette PB peut être mise en réserve dans une *Provision pour Participation aux Bénéfices* (PPB) que nous appellerons *Provision pour Participations aux Excedents* (PPE) dans la suite de ce mémoire. La PPE doit être reversée aux assurés dans un délai de huit ans au maximum.

Les **unités de compte (UC)** offrent une exposition directe aux marchés financiers. Contrairement au fonds en euros, le risque d'investissement est entièrement porté par l'assuré. L'assureur ne garantit pas la valeur du capital, mais un nombre de parts d'actifs (OPCVM, actions, SCPI, etc.). La valeur de

l'épargne fluctue ainsi au gré des marchés, offrant un potentiel de rendement supérieur à long terme, mais exposant également à un risque de perte en capital. Pour l'assureur, ce support est plus simple à gérer car il n'implique pas de garanties financières particulières.

Ces supports sont proposés via deux grandes familles de contrats. Les contrats monosupports permettent d'investir sur un seul type de fonds (soit en euros, soit en UC). Les contrats multisupports sont les plus répandus, quant à eux, combinent au moins un fonds en euros et plusieurs supports en unités de compte, permettant à l'épargnant de répartir son investissement selon son profil de risque. Dans le cadre de cette étude, le portefeuille analysé se compose de contrats multisupports avec une répartition représentative du marché français, soit approximativement 60% en fonds euros et 40% en unités de compte.

	Ensemble		Euros		Unités de compte ²	
	En Md€	Var.	En Md€	Var.	En Md€	Var.
Cotisations	174,9	+14,7 %	108,6	+19,2 %	66,3	+8,1 %
Prestations	146,4	-3,1 %	113,3	-5,5 %	33,1	+6,4 %
Collecte nette	+28,5	+27,1 Md€	-4,7	+24,1 Md€	+33,2	+3,0 Md€
Encours	+1 985,8	+3,9 %	+1 398,6	+1,4 %	+587,1	+10,3 %
dont provisions mathématiques	+1932,2	+4,4 %	+1345,0	+2,0 %	+587,1	+10,3 %
dont provision pour participation aux bénéfices	+53,6	-11,1 %	+53,6	-11,1 %	n.a.	n.a.

FIGURE 1.2 – Le marché de l'assurance vie en France en 2024 (graphique temporel)

Ces contrats sont soumis à un ensemble de réglementations prudentielles visant à garantir la solvabilité des assureurs et la protection des assurés. La directive Solvabilité II encadre ces exigences à l'échelle européenne. La partie suivante détaille le cadre réglementaire de Solvabilité II, ses implications pour les assureurs vie, ainsi que les outils de modélisation stochastique utilisés pour répondre à ces exigences.

1.2 Le cadre prudentiel Solvabilité II

Entrée en vigueur le 1er janvier 2016, la directive Solvabilité II régit le cadre prudentielle pour la quasi-totalité des assureurs et réassureurs de l'Union Européenne. Son principal objectif est d'harmoniser les pratiques du secteur, d'assurer une protection optimale des assurés et de garantir que les compagnies puissent honorer leurs engagements en toutes circonstances. Pour ce faire, elle instaure une approche économique et prospective, fondée sur une évaluation fine des risques et structurée en trois piliers inter-dépendants. Seuls les deux premiers piliers seront présentés car ils sont plus pertinents dans le cadre de ce mémoire.

1.2.1 Le Pilier 1 : Exigences quantitatives et Bilan Prudentiel

Le Pilier 1 définit les exigences quantitatives, au cœur desquelles se trouve le Bilan Prudentiel. Il s'agit d'une vision économique du bilan comptable où les actifs et les passifs sont évalués de manière cohérente avec leur valeur de marché (*market-consistent*). La structure de cette section suivra l'équation fondamentale du bilan prudentiel :

$$\text{Actifs} = \text{Provisions Techniques} + \text{Fonds Propres} \quad (1.1)$$

Chaque terme de cette équation sera détaillé successivement.

BILAN ECONOMIQUE SOUS SOLVABILITE II		
Actifs	Fonds propres	Capital excédentaire
		SCR
		MCR
	Provisions techniques	Risk Margin
		Best Estimate

FIGURE 1.3 – Bilan économique sous Solvabilité II (graphique temporaire)

1.2.1.1 L'évaluation des Actifs à la Valeur de Marché

Le premier terme du bilan, Actifs, sont comptabilisés à leur Valeur de Marché (VM). Cette approche vise à refléter la valeur la plus juste et actuelle des ressources dont dispose l'assureur pour couvrir ses engagements.

1.2.1.2 Les Provisions Techniques : Cœur de l'évaluation du passif

Les Provisions Techniques (PT) représentent la valeur des engagements de l'assureur envers ses assurés. Elles se décomposent en deux parties : le *Best Estimate* (BE) et la Marge de Risque (*Risk Margin* - RM).

$$PT = \text{Best Estimate (BE)} + \text{Marge de Risque (RM)} \quad (1.2)$$

La valorisation des provisions techniques, au cœur du bilan prudentiel, ne peut se contenter d'une vision unique et figée du futur. La présence d'options et de garanties dans les contrats d'épargne impose de distinguer deux approches complémentaires :

- **L'approche déterministe** est un outil de pilotage. Elle repose sur une projection unique des variables économiques. Bien qu'insuffisante pour la valorisation prudentielle, elle demeure un outil fondamental pour l'élaboration du *business plan* et la communication d'un scénario central. Sa limite principale est son incapacité à valoriser les risques asymétriques.
- **L'approche stochastique** est un outil de valorisation. Elle explore un grand nombre de futurs possibles à l'aide d'un **Générateur de Scénarios Économiques (GSE)**. Cet outil produit des milliers de simulations cohérentes des marchés financiers. La valeur d'un indicateur est alors obtenue en calculant la moyenne des résultats sur l'ensemble de ces scénarios (méthode de Monte-Carlo). Cette exploration est indispensable pour quantifier le coût réel des garanties optionnelles (Taux Minimum Garanti, etc.).

La différence de valeur entre ces deux approches est capturée par le concept de **TVOG (*Time Value of Options and Guarantees*)**. En imposant une approche stochastique, Solvabilité II assure une valorisation *market-consistent* des engagements.

Le **Best Estimate (BE)**, ou *Best Estimate Liability (BEL)*, représente la meilleure estimation de la valeur actuelle des flux de trésorerie futurs liés aux engagements d'assurance. Son calcul est réalisé sur un horizon long (40-60 ans) en *run-off* (portefeuille en extinction, pas de nouvelles souscriptions). Il est obtenu par la moyenne des flux actualisés sur un grand nombre de simulations économiques stochastiques en univers risque neutre :

$$BEL = \mathbb{E}^{\mathbb{Q}} \left[\sum_{j=1}^T CF(j) \cdot e^{-\int_0^j r(s) ds} \right] \approx \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^T \frac{CF_i(j)}{(1 + r_{i,j})^j} \quad (1.3)$$

Où N est le nombre de simulations, T l'horizon de projection, $CF_i(j)$ le flux de trésorerie net de l'année j pour la simulation i , et $r_{i,j}$ le taux d'actualisation sans risque pertinent.

La **Marge de Risque (RM)** complète le Best Estimate. Elle correspond à la rémunération du capital réglementaire qui doit être immobilisé pour couvrir les risques non-financiers (longévité, rachat, etc.) jusqu'à l'extinction du portefeuille. Son calcul repose sur une approche dite de "Coût du Capital" (*Cost*

of Capital - CoC), qui consiste à actualiser le coût futur de détention de ce capital :

$$RM = CoC_{rate} \times \sum_{j=0}^T \frac{SCR_{non-fi}(j)}{(1 + r_{j+1})^{j+1}} \quad (1.4)$$

Où CoC_{rate} est le coût du capital (fixé à 6%), et $SCR_{non-fi}(j)$ est la part du SCR couvrant les risques non-financiers à l'année j .

1.2.1.3 Les Fonds Propres et les Exigences de Capital

Les **Fonds Propres Prudentiels**, aussi appelés **NAV (Net Asset Value)**, constituent la richesse de l'assureur. Ils sont définis par la différence entre la valeur des actifs et celle des engagements :

$$NAV = VM_{Actifs} - (BE + RM) \quad (1.5)$$

Ce sont ces fonds propres qui doivent permettre à l'assureur d'absorber des pertes inattendues. Solvabilité II définit donc deux niveaux d'exigence de capital portant sur la NAV.

Le **Solvency Capital Requirement (SCR)** est le montant de fonds propres nécessaire pour absorber des pertes exceptionnelles. Il est calibré pour correspondre à la **Value-at-Risk (VaR) à 99,5%** de la NAV sur un horizon d'un an. Autrement dit, il s'agit du capital qui doit permettre à l'assureur de faire face à ses engagements sur l'année à venir avec une probabilité de 99,5%. En cas de non-respect, l'assureur fait l'objet d'un suivi renforcé par le régulateur.

Le calcul du SCR peut se faire via un modèle interne (spécifique à l'assureur) ou, plus communément, via la **Formule Standard** définie par la réglementation. Cette dernière est une approche modulaire qui décompose le risque total en plusieurs modules et sous-modules (risque de marché, de souscription, de contrepartie, etc.).

Pour chaque risque élémentaire x , le capital requis est calculé comme la perte de NAV consécutive à un choc instantané et calibré sur ce risque :

$$SCR_x = \Delta NAV = NAV_{central} - NAV_{choc\ x} \quad (1.6)$$

Les SCR des différents modules sont ensuite agrégés en prenant compte des corrélations prédéfinies entre les risques. L'agrégation de deux modules de risque i et j se fait via la formule :

$$SCR_{i,j} = \sqrt{SCR_i^2 + SCR_j^2 + 2 \times Corr_{i,j} \times SCR_i \times SCR_j} \quad (1.7)$$

où $Corr_{i,j}$ est le coefficient de corrélation entre les risques i et j fourni par la réglementation. Cette agrégation est appliquée de manière hiérarchique pour obtenir le SCR total, appelé **Basic Solvency Capital Requirement (BSCR)** :

$$SCR_{total} = \sqrt{\sum_i \sum_j SCR_i \times SCR_j \times Corr_{i,j}} \quad (1.8)$$

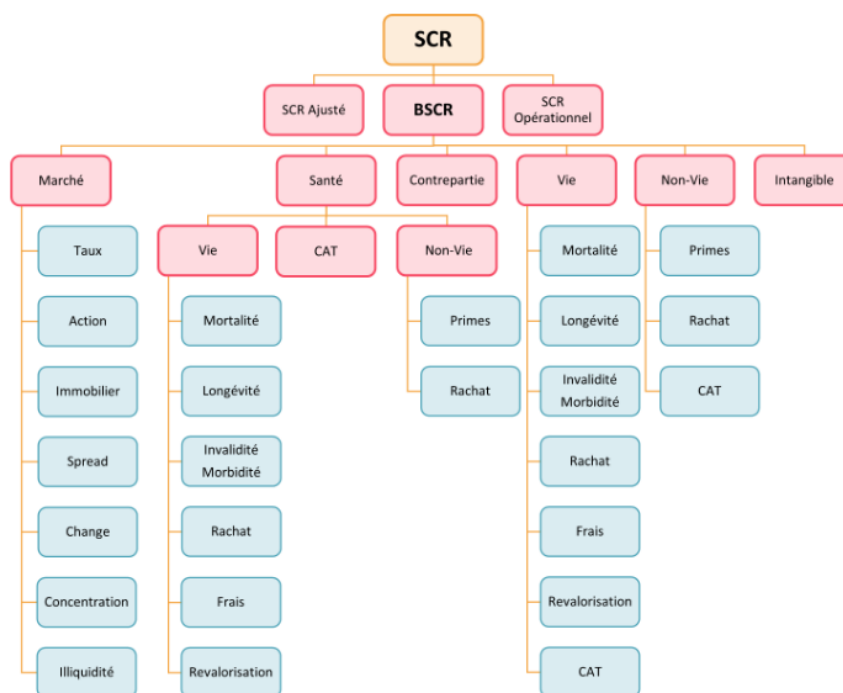


FIGURE 1.4 – Schéma des modules et sous-modules du SCR en Formule Standard

1.2.2 Le Pilier 2 : Exigences qualitatives et gouvernance

Ce pilier se concentre sur la supervision, la gestion des risques et la gouvernance interne. Il impose aux assureurs de mettre en place un système de gouvernance sain, prudent et proportionné. Cela inclut une structure organisationnelle transparente, des politiques écrites claires, et un système de contrôle interne robuste. La direction doit être assurée par au moins deux dirigeants effectifs (principe des 4 yeux) qui doivent répondre à des exigences de compétence et d'honorabilité (*fit and proper*).

Ce système s'articule autour de quatre fonctions clés indépendantes : la fonction actuarielle, la gestion des risques, l'audit interne et la conformité.

L'élément central du Pilier 2 est l'**ORSA** (*Own Risk and Solvency Assessment*). Il s'agit d'un processus interne par lequel l'assureur évalue, sur un horizon de 3 à 5 ans, l'adéquation entre son profil de risque spécifique, ses limites de tolérance et ses besoins globaux en solvabilité. C'est un outil de pilotage stratégique qui permet d'aller au-delà des hypothèses standards pour refléter la stratégie propre de l'entreprise.

1.3 Les Générateurs de Scénarios Économiques (GSE)

Le Générateur de Scénarios Économiques (GSE) est un outil mathématique central dans la modélisation stochastique. Il a pour fonction de simuler de multiples trajectoires futures pour les principales variables financières (taux d'intérêt, performance des actions, inflation, etc.). La qualité des projections ALM dépendant directement de la robustesse du GSE, il est nécessaire de distinguer deux cadres de modélisation qui coexistent.

Bien que ces deux univers soient complémentaires, la réglementation Solvabilité II assigne à chacun un rôle très précis pour le calcul des différents indicateurs prudentiels. Le tableau suivant synthétise cette répartition des tâches.

TABLE 1.1 – Répartition des calculs Solvabilité II par univers de projection

Univers Risque Neutre (Q)	Univers Monde Réel (P)
Indicateurs du Pilier 1 : <ul style="list-style-type: none"> — Best Estimate Liability (BEL) — Marge de Risque (RM) — Solvency Capital Requirement (SCR) — Bilan Prudentiel et NAV 	Exercices du Pilier 2 : <ul style="list-style-type: none"> — ORSA (Own Risk and Solvency Assessment) — Business Plan et planification stratégique — Test de la pérennité du modèle
Exercices du Pilier 2 : <ul style="list-style-type: none"> — ORSA (Own Risk and Solvency Assessment) 	
Finalité : Valorisation <i>Market-Consistent</i> à un instant t.	Finalité : Pilotage stratégique et prospectif.

La distinction entre ces deux approches est donc fondamentale : l'une sert à valoriser, l'autre à piloter. Les sections suivantes détaillent les modèles mathématiques sous-jacents à chaque univers.

1.3.1 L'univers Risque Neutre (Q) : un cadre pour la valorisation

L'univers **Risque Neutre (Q)** est un cadre de valorisation théorique, requis par Solvabilité II pour les calculs *market-consistent*. Son objectif n'est pas de prédire l'évolution réelle des marchés, mais de calculer la valeur risque neutralisée d'un actif ou d'un passif à la date de calcul, en se fondant sur les prix de marché observés. Dans cet univers, on postule que tous les investisseurs sont indifférents au risque, ce qui implique que le rendement espéré de n'importe quel actif est égal au taux d'intérêt sans risque. Cette construction, fondée sur l'absence d'opportunité d'arbitrage, est indispensable pour valoriser de manière cohérente les options et garanties complexes des contrats d'assurance. La valeur V_0 d'un flux de trésorerie futur CF_T est alors son espérance mathématique sous cette probabilité risque neutre, actualisée au taux sans risque $r(s)$:

$$V_0 = \mathbb{E}^Q \left[CF_T \cdot e^{-\int_0^T r(s) ds} \right] \quad (1.9)$$

Cet univers constitue le fondement du Pilier 1 de Solvabilité II, utilisé pour le calcul du Best Estimate Liability (BEL) et du Solvency Capital Requirement (SCR).

1.3.1.1 Modélisation des taux d'intérêt : le modèle de Hull & White

Pour les taux d'intérêt, le modèle de **Hull & White à un facteur** est une référence dans le cadre réglementaire. Son principal avantage est sa capacité à se calibrer parfaitement à la courbe des taux sans risque initiale, telle que fournie par l'EIOPA.

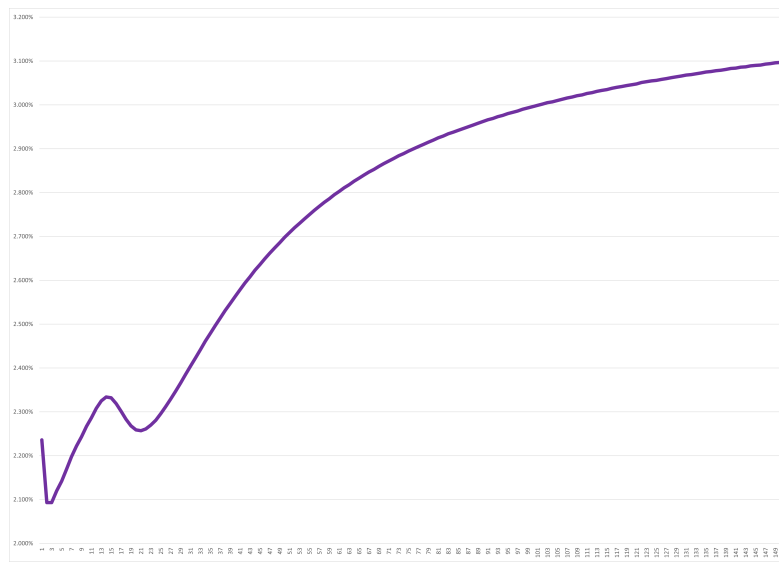


FIGURE 1.5 – Courbe des taux sans risque sans *Volatility Adjustment* au 31/12/2024 publiée par l'EIOPA

Cette flexibilité est obtenue grâce à un paramètre de retour à la moyenne $\theta(t)$ qui dépend du temps. Son équation différentielle stochastique (EDS) s'écrit :

$$dr_t = (\theta(t) - ar_t)dt + \sigma dW_t^{\mathbb{Q}} \quad (1.10)$$

où r_t est le taux d'intérêt court, a la vitesse de retour à la moyenne, σ la volatilité et $W_t^{\mathbb{Q}}$ un mouvement brownien sous la mesure risque neutre. Pour des raisons de calcul, nous utilisons la solution discrète de cette EDS :

$$r_{t+h} = r_t e^{-ah} + \theta(t+h) - \theta(t)e^{-ah} + \sigma \sqrt{\frac{1 - e^{-2ah}}{2a}} Z \quad (1.11)$$

1.3.1.2 Modélisation des actions et de l'immobilier : le modèle de Black & Scholes

Pour les actifs risqués comme les actions ou l'immobilier, le modèle de **Black & Scholes** est couramment utilisé. Conformément à la logique risque neutre, le rendement espéré (la dérive du processus) est le taux sans risque r_t . L'EDS du prix de l'actif S_t est :

$$dS_t = r_t S_t dt + \sigma S_t dW_t^{\mathbb{Q}} \quad (1.12)$$

où S_t est le prix de l'actif, r_t le taux sans risque et σ la volatilité de l'actif. La solution de cette EDS est donnée par :

$$S_t = S_0 \exp \left(\int_0^t \left(r_s - \frac{\sigma^2}{2} \right) ds + \int_0^t \sigma dW_s^{\mathbb{Q}} \right) \quad (1.13)$$

En pratique, on utilise sa solution discrétisée pour simuler les trajectoires de prix sur un pas de temps h :

$$S_{t+h} = S_t \exp \left(\left(r_t - \frac{\sigma^2}{2} \right) h + \sigma \sqrt{h} Z \right) \quad (1.14)$$

où Z est une variable aléatoire suivant une loi normale centrée réduite $\mathcal{N}(0, 1)$.

1.3.2 L'univers Monde Réel (\mathbb{P}) : un outil de pilotage stratégique

À l'inverse de l'univers risque neutre, l'**univers Monde Réel (\mathbb{P})** vise à générer des scénarios réalistes pour refléter une évolution plausible des marchés. Son objectif est la projection et la planification stratégique, notamment pour l'exercice ORSA (Pilier 2).

La différence fondamentale réside dans l'introduction d'une **prime de risque** pour rémunérer la volatilité supportée par les investisseurs. Le rendement espéré d'un actif risqué est donc supérieur au taux sans risque, calibré sur des données historiques et des anticipations d'experts :

$$\mathbb{E}^{\mathbb{P}}[\text{Rendement de l'actif}] = \text{Taux sans risque} + \text{Prime de risque} \quad (1.15)$$

Les modèles utilisés, bien que similaires dans leur forme à ceux de l'univers \mathbb{Q} (par exemple, Vasicek pour les taux ou Black & Scholes pour les actions), sont modifiés pour intégrer cette prime. La dérive du processus stochastique n'est plus le taux sans risque r_t , mais un rendement espéré monde réel μ . En somme, si l'univers \mathbb{Q} valorise les engagements à un instant t , l'univers \mathbb{P} permet d'exprimer la situation financière de l'entreprise dans le futur, ce qui le rend indispensable pour le pilotage stratégique.

1.3.3 Synthèse des deux univers

Le tableau suivant résume les caractéristiques et les usages des deux univers de projection. Si l'univers risque neutre \mathbb{Q} répond à la question « *Combien vaut cet engagement aujourd'hui ?* », l'univers monde réel \mathbb{P} répond à « *Quelle sera ma situation financière demain ?* ».

TABLE 1.2 – Synthèse comparative des univers de projection

Critère	Univers Risque Neutre (\mathbb{Q})	Univers Monde Réel (\mathbb{P})
Objectif	Valorisation <i>Market-Consistent</i> (Pilier 1 : BEL, SCR). Calculer une valeur juste à $t = 0$.	Projection stratégique (Pilier 2 : ORSA, Business Plan). Simuler des futurs plausibles.
Rendement Espéré (Actifs risqués)	Taux sans risque (r_t). Aucune prime de risque.	Taux sans risque + Prime de risque ($\mu = r + \text{prime}$).
Modèle de Taux Typique	Hull & White . Flexible, calibré à la courbe des taux initiale.	Vasicek . Économique, retour à une moyenne de long terme.
Calibration	Calibré sur les prix des instruments financiers actuels (courbe des taux, volatilités implicites).	Calibré sur des données historiques et des anticipations d'experts (primes de risque).

Pour la suite de ce mémoire, je vais me concentrer sur l'univers risque neutre \mathbb{Q} , car il est le plus pertinent pour les calculs prudentiels et la gestion Actif-Passif dans le cadre des calculs liés au pilier 1 de Solvabilité II.

1.4 La représentation du passif : le concept de *Model Point*

1.4.1 La nécessité de l'agrégation

Les portefeuilles d'assurance vie comptent souvent des centaines de milliers, voire des millions de polices. Une modélisation "police à police" est techniquement possible mais informatiquement très chronophage voire irréalisable pour des calculs stochastiques complexes comme ceux requis par les modèles ALM. La charge de calcul deviendrait prohibitive. La simplification du portefeuille de passif n'est donc pas un choix, mais une contrainte opérationnelle majeure.

La réponse standard à cette contrainte est la création de *Model Points* (MP). Un MP est un contrat synthétique représentant un agrégat de polices partageant des caractéristiques homogènes. L'objectif est de réduire drastiquement le volume de données à traiter tout en préservant les propriétés actuarielles et financières essentielles du portefeuille complet. La qualité de la représentation dépend directement de la pertinence des critères de regroupement (caractéristiques du produit, de l'assuré, du contrat), souvent optimisés par des techniques de classification statistique (*clustering*).

1.4.2 Les impacts de l'agrégation en MP sur les indicateurs S2

L'utilisation des *Model Points* constitue la méthode standard pour agréger le passif et rendre les modèles ALM opérationnels. Cependant, la manière dont ces groupes de contrats sont formés à partir des polices individuelles a un impact direct et significatif sur la valorisation des indicateurs réglementaires et leur sensibilité aux chocs.

La problématique centrale de ce mémoire est donc d'optimiser cette étape fondamentale de l'agrégation. En partant du portefeuille granulaire "police à police", nous cherchons à définir, comparer et tester différentes méthodes de regroupement pour créer des *Model Points*. L'objectif est de construire une méthodologie de simplification optimale, c'est-à-dire celle qui minimise l'erreur d'agrégation tout en garantissant la plus grande stabilité des indicateurs Solvabilité II (SCR, Marge de Risque, NAV) lors du calcul des différentes sensibilités. Cette section introductive pose donc les fondations de notre analyse : la simplification étant une nécessité, comment s'assurer que la méthode de regroupement choisie est la plus fidèle et la plus robuste possible ?

Chapitre 2

La Gestion Actif-Passif et le Modèle de Simulation en Python

2.1 Principes et Enjeux de la Gestion Actif-Passif (ALM)

La gestion Actif-Passif (ALM) trouve son origine dans une particularité fondamentale du secteur de l'assurance énoncé précédemment dans le mémoire : le **cycle de production inversé**. Contrairement à une entreprise classique qui vend un produit avant d'en percevoir le revenu, un assureur collecte des primes aujourd'hui en échange de la promesse de verser des prestations dans un futur lointain et incertain. Ce décalage temporel est au cœur du modèle économique de l'assurance vie.

Ce mécanisme engendre une inadéquation structurelle (*mismatch*) entre les deux côtés du bilan. D'une part, le passif est constitué d'engagements de longue durée, dont l'échéance et le montant sont soumis à des aléas (mortalité, comportement de rachat des assurés). D'autre part, pour couvrir ces engagements, l'assureur investit les primes sur les marchés financiers, constituant un actif dont la valeur et les flux sont, par nature, volatiles et dépendants du contexte économique.

Cette inadéquation est renforcée par une interdépendance dynamique et complexe entre l'actif et le passif.

- **Le passif influe sur l'actif** : Le versement des prestations (rachats, décès) contraint l'assureur à liquider une partie de ses actifs, parfois dans des conditions de marché défavorables.
- **L'actif influe sur le passif** : La performance des actifs financiers a un impact direct sur le niveau des engagements. C'est notamment le cas de la **Participation aux Bénéfices (PB)**, qui dépend des résultats financiers générés par l'assureur. Le Code des assurances impose une redistribution minimale aux assurés, calculée comme suit :

$$PB_{\min\text{Reg}} = 85\% \times \max(\text{RésFi}, 0) + \begin{cases} 90\% \times \text{RésTech} & \text{si } \text{RésTech} \geq 0 \\ 100\% \times \text{RésTech} & \text{si } \text{RésTech} < 0 \end{cases} \quad (2.1)$$

Où le Résultat Financier (RésFi) est directement issu de la performance des actifs et le Résultat Technique (RésTech) des risques de mortalité et de rachat.

La gestion Actif-Passif est donc la discipline qui vise à piloter les risques nés de cette interdépendance afin d'assurer la solvabilité et d'optimiser la rentabilité de l'acteur.

2.1.1 La Modélisation ALM : un Outil de Projection Essentiel

Pour quantifier et piloter les risques complexes découlant de l'inadéquation actif-passif, les assureurs ont recours à des modèles de projection actuariels sophistiqués, communément appelés modèles de Gestion Actif-Passif ou modèles ALM. L'utilité principale de ces modèles est de projeter le bilan d'un assureur, qu'il s'agisse d'un bilan prudentiel sous Solvabilité 2 ou d'un bilan comptable sous les normes IFRS17, afin d'évaluer la santé financière future de l'organisme sur un horizon de long terme. Leur fonctionnement repose sur la combinaison de deux piliers fondamentaux : des scénarios prospectifs sur l'environnement économique et financier, et des hypothèses sur le comportement futur des assurés (lois de mortalité, de rachat, etc.).

L'approche de modélisation peut être déterministe ou stochastique, chaque approche répondant à des objectifs d'analyse distincts.

L'approche **déterministe** consiste à projeter le bilan de l'assureur selon une trajectoire unique et prédéfinie de l'environnement économique. Cette trajectoire, qualifiée de **scénario central**, est généralement construite à partir de la courbe des taux sans risque fournie par l'EIOPA. Elle permet d'obtenir le (*Best Estimate*) central, c'est à dire la somme des flux futurs actualisés du portefeuille dans un contexte économique considéré comme le plus probable, servant de base pour le plan d'affaires et la valorisation prudentielle.

Cependant, une approche déterministe ne peut à elle seule capturer l'éventail des risques, notamment ceux liés aux options et garanties financières (par exemple, les Taux Minimum Garantis ou les options de rachat) dont le coût ne se matérialise que dans des conditions de marché adverses. Pour pallier cette limite, une approche **stochastique** est nécessaire. Celle-ci s'appuie sur un **Générateur de Scénarios Économiques (GSE)** pour simuler un grand nombre (souvent plusieurs milliers) de trajectoires économiques futures possibles, chacune représentant une évolution plausible des marchés financiers.

Chaque scénario économique généré sert alors d'input pour une projection complète du modèle ALM. En agrégeant les résultats de ces multiples projections via la **méthode de Monte Carlo**, l'assureur obtient non plus une seule valeur, mais une distribution des résultats possibles. L'objectif final est de disposer, pour chaque trajectoire, du détail des flux financiers à la maille la plus fine. Cette granularité permet une analyse statistique approfondie des risques, comme le calcul de quantiles (Value at Risk à 99.5%) pour déterminer le capital de solvabilité requis (SCR) sous Solvabilité 2. Ces modèles stochastiques sont donc au cœur de l'évaluation des risques et de la prise de décision stratégique, et constituent le fondement du modèle de simulation qui sera détaillé dans la suite de ce chapitre.

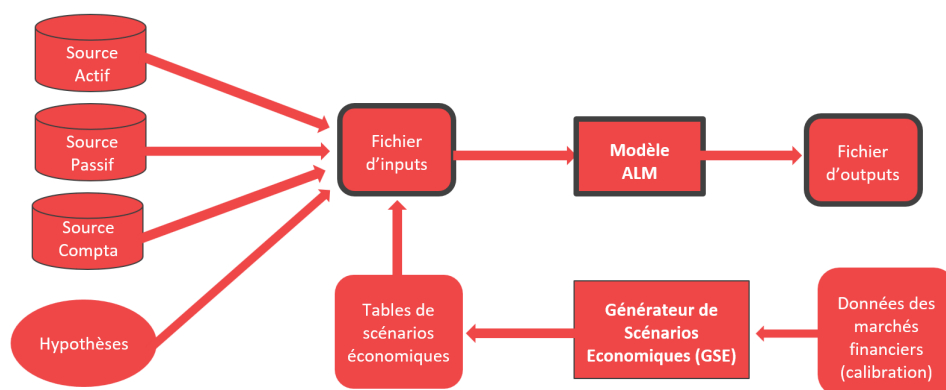


FIGURE 2.1 – Fonctionnement d'un modèle ALM (graphique temporel)

2.2 Architecture et Fonctionnement du Modèle de Projection

L'objectif de cette section est de détailler l'architecture et le séquençement des opérations du modèle ALM développé pour les besoins de cette étude. Le modèle a été conçu pour simuler de manière dynamique et séquentielle le bilan d'un assureur vie sur un horizon de projection pluriannuel.

Son fonctionnement global peut être décomposé en trois phases principales, comme illustré dans la figure 2.2 :

1. **Phase d'Initialisation** : Préparation et validation des données d'entrée, et application des choix réglementaires à la date de départ.
2. **Boucle de Projection Annuelle** : Cœur du modèle qui simule, année après année, l'évolution du bilan selon une séquence d'événements prédéfinis.
3. **Phase de Finalisation** : Calcul des indicateurs prudentiels et génération des résultats en fin de projection.

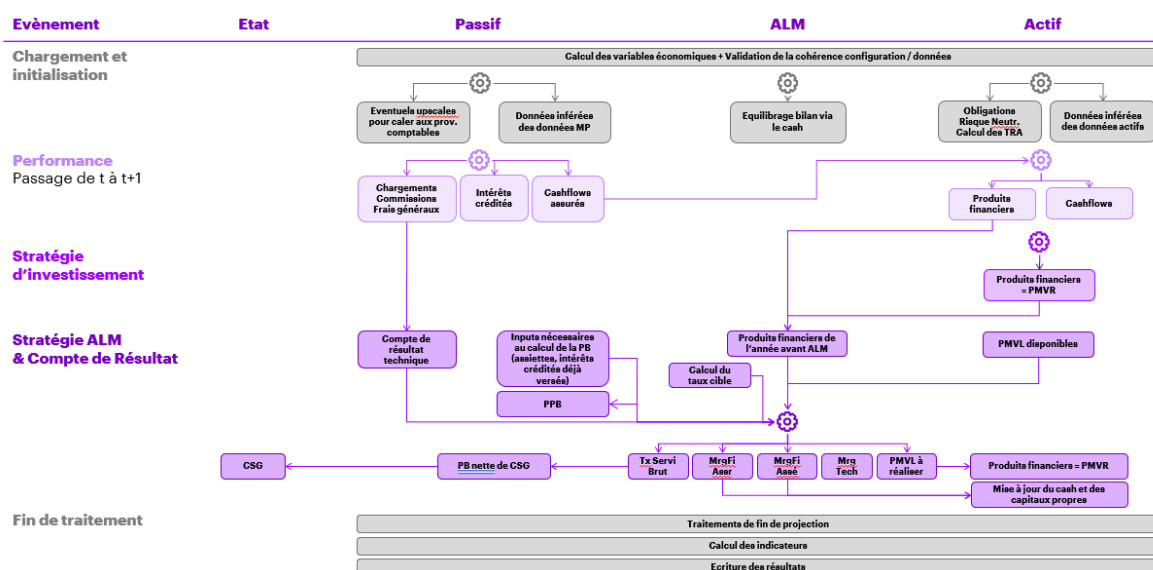


FIGURE 2.2 – Architecture générale et séquençement des événements du modèle ALM. (graphique temporaire)

2.2.1 Phase 1 : Initialisation du Modèle

Cette première phase prépare l'environnement de projection. Elle est elle-même divisée en trois sous-étapes.

2.2.1.1 Chargement et Validation des Données d'Entrée

Le modèle est alimenté par un ensemble de données exhaustif, regroupées en quatre catégories :

- **Données Économiques et Financières** : Issues du Générateur de Scénarios Économiques (GSE), elles comprennent les courbes de taux, les taux d'inflation et les performances des différentes classes d'actifs pour chaque scénario stochastique.
- **Portefeuille d'Actifs** : Les *model points* d'actifs représentant l'ensemble des placements de l'assureur (obligations à taux fixe et variable, actions et immobilier).
- **Portefeuille de Passifs** : Les *model points* d'épargne décrivant les engagements envers les assurés.
- **Hypothèses de Modélisation** : Un ensemble de tables paramétrant le comportement futur (stratégie d'investissement, stratégie ALM, règles de participation aux bénéfices, tables de mortalité, de rachat) et les chocs prudentiels Solvabilité 2.

Une étape de validation est systématiquement réalisée pour assurer la cohérence et la qualité des données chargées. Par exemple, nous vérifions qu'il n'y a pas de Provisions mathématiques négatives ou que la somme des taux d'affectation des différentes classes d'actifs est bien égale à 1.

2.2.1.2 Application des Chocs Solvabilité 2 en $t=0$

L'application des chocs instantanés en $t=0$ est une étape fondamentale du calcul du SCR (Solvency Capital Requirement) selon la formule standard de Solvabilité 2. L'objectif est de mesurer la résilience de l'assureur face à une série de scénarios de crise prédéfinis, calibrés pour représenter un événement se produisant une fois tous les 200 ans.

Pour chaque choc, le modèle crée un "axe analytique" : les données d'entrée sont dupliquées, puis les paramètres pertinents sont modifiés conformément au scénario de choc. La différence entre la valeur des fonds propres dans le scénario central (sans choc) et leur valeur dans le scénario choqué détermine le besoin en capital pour ce risque spécifique.

Les principaux chocs, ou modules de risque, se répartissent en plusieurs catégories :

- **Le risque de marché** : Il regroupe les risques liés aux fluctuations des marchés financiers.
 - **Choc de taux d'intérêt** : Simule une variation soudaine, à la hausse ou à la baisse, de la courbe des taux d'intérêt. Ce choc a un impact majeur sur la valeur des actifs (notamment les obligations) et des passifs (qui sont actualisés avec cette courbe des taux).
 - **Choc actions** : Modélise une baisse brutale des marchés actions. La formule standard distingue généralement deux types de chocs actions (Type 1 et Type 2) en fonction de la nature et de la diversification des investissements. Par exemple, un choc de -33.84% pourrait s'appliquer à un portefeuille d'actions diversifié.

- **Choc immobilier** : Représente une chute de la valeur du marché immobilier. Le paramètre de -25% est une valeur standard pour ce type de risque.
- **Choc de spread** : Concerne le risque de crédit sur les obligations et les prêts. Il simule un élargissement des spreads de crédit, ce qui diminue la valeur de marché de ces actifs.
- **Le risque de souscription vie** : Il est lié aux aléas inhérents à l'activité d'assurance vie.
 - **Choc de mortalité** : Simule une augmentation soudaine et permanente du taux de mortalité (par exemple, +15%). Ce choc est particulièrement impactant pour les contrats de prévoyance où l'assureur doit verser un capital en cas de décès.
 - **Choc de longévité** : À l'inverse, il modélise une baisse permanente du taux de mortalité (les assurés vivent plus longtemps que prévu). Ce risque affecte principalement les rentes viagères, pour lesquelles l'assureur doit verser des prestations plus longtemps.
 - **Choc de rachat** : Simule une variation massive et soudaine du comportement des assurés en matière de rachat de leurs contrats. Il se décline en trois sous-modules : une hausse des taux de rachat, une baisse, et un scénario de rachat de masse (catastrophique).
 - **Choc de dépenses** : Modélise une augmentation imprévue des frais de gestion de l'assureur, combinée à une hausse de l'inflation de ces frais.
 - **Choc de catastrophe** : Concerne les événements extrêmes, comme une pandémie, entraînant une augmentation brutale et temporaire de la mortalité (par exemple, une hausse de 0.15 points de pourcentage du taux de mortalité sur une année).

2.2.1.3 Préparation à la Boucle de Projection

Enfin, les tables de données sont préparées pour la projection en y ajoutant les dimensions d'analyse fondamentales : le **scénario économique**, la **période de projection** (l'année) et l'**événement** intra-annuel.

2.2.2 Phase 2 : La Boucle de Projection Annuelle

Le cœur du modèle est une boucle itérative qui projette le bilan année par année. Pour chaque pas de temps, trois événements clés sont modélisés séquentiellement.

2.2.2.1 Événement 1 : Performance

Cette première étape simule le "passage du temps" sans intervention active du management.

- **Côté Passif** : Les engagements évoluent sous l'effet des flux biométriques (décès), comportements (rachats), et contractuels (arrérages de rente). La provision mathématique est revalorisée en

appliquant les Taux Minimums Garantis (TMG) pour les fonds euros et la performance des marchés pour les Unités de Compte (UC).

- **Côté Actif** : La valeur des actifs est mise à jour pour refléter la performance des marchés. Les revenus (coupons, dividendes) sont encaissés et alimentent la trésorerie.

2.2.2.2 Événement 2 : Stratégie d'Investissement

Cette étape modélise les décisions de gestion financière. L'algorithme simule la politique d'investissement en cherchant à maintenir une allocation d'actifs cible. En fonction de la trésorerie disponible, le modèle arbitre entre les différentes classes d'actifs, générant des ordres d'achats et de ventes.

2.2.2.3 Événement 3 : Stratégie ALM et Clôture du Compte de Résultat

C'est l'étape finale de l'exercice annuel, qui vise à équilibrer le bilan.

1. **Détermination du Résultat Financier** : Le modèle agrège l'ensemble des produits financiers générés.
2. **Application de la Stratégie de Participation aux Bénéfices (PB)** : Le modèle calcule la PB à distribuer, en respectant les contraintes réglementaires et en visant un taux cible. Il peut activer des leviers comme la reprise de la Provision pour Participation aux Excédents (PPE) ou la réalisation de plus-values latentes.
3. **Clôture des Comptes** : Le résultat technique et financier sont finalisés pour établir le compte de résultat et s'assurer de l'équilibre du bilan de clôture.

2.2.3 Phase 3 : Finalisation et Génération des Outputs

Une fois la boucle achevée pour toutes les années et tous les scénarios, le modèle entre dans sa phase finale.

2.2.3.1 Calcul des Indicateurs Prudentiels

Le principal objectif est de calculer le **Best Estimate (BE)** des passifs. Pour chaque scénario, le modèle agrège les flux de passifs futurs et les actualise en utilisant les courbes de taux sans risque correspondantes. La moyenne des BE sur l'ensemble des scénarios fournit la valeur centrale, tandis que la distribution permet de calculer le capital de solvabilité requis (SCR).

2.2.3.2 Génération des Rapports de Sortie

Enfin, le modèle produit un ensemble de rapports détaillés (similaires aux QRTs réglementaires) présentant les bilans projetés, les comptes de résultat, la décomposition du BE et du SCR, et d'autres indicateurs clés pour l'analyse actuarielle.

2.3 Limites Actuelles du modèle

A date, les fonds propres ne sont pas modélisés dans le modèle ALM. En effet, le modèle se concentre sur la projection de l'actif et du passif, ainsi que sur les interactions entre les deux, mais n'intègre pas encore la dynamique des fonds propres. Cette limitation est importante car les fonds propres jouent un rôle crucial dans la solvabilité et la résilience financière de l'assureur. Leur modélisation permettrait de mieux évaluer l'impact des stratégies de gestion sur la solidité financière globale de l'entreprise. L'absence des fonds propres nous empêche de calculer des indicateurs cruciaux tels que le ratio de solvabilité.

Chapitre 3

Construction d'un Générateur de Portefeuilles de Passifs Réaliste

3.1 Objectifs Stratégiques et Contraintes Techniques

La capacité à tester la robustesse des modèles et la pertinence des analyses de sensibilité repose sur un prérequis fondamental : la disponibilité de données de passif variées et réalistes. Pour un cabinet de conseil, où l'accès aux portefeuilles des clients n'est pas systématique, la faculté de générer des portefeuilles synthétiques, mais représentatifs du marché, constitue un atout stratégique majeur. C'est dans ce contexte qu'un générateur de portefeuilles de passifs a été conçu et développé dans le cadre de ce mémoire.

Ce chapitre a pour vocation de présenter cet outil essentiel. Nous détaillerons les besoins stratégiques et analytiques auxquels il répond, la méthodologie de génération retenue, les contraintes techniques rencontrées et les données qui ont été utilisées pour rendre le portefeuille le plus réaliste possible.

3.1.1 Définition du générateur de portefeuilles de passifs

Un générateur de portefeuille de passifs est un outil conçu pour créer, de manière algorithmique, des ensembles de données synthétiques qui imitent avec réalisme des portefeuilles de contrats d'assurance-vie. Plutôt que de s'appuyer sur des données réelles, souvent confidentielles ou indisponibles, cet outil simule les caractéristiques fondamentales des assurés (âge, sexe, etc.) et de leurs contrats (type de produit, montant de la provision mathématique, date de souscription, etc.).

L'objectif n'est pas de produire des données aléatoires, mais de générer un portefeuille dont les propriétés statistiques — distributions, corrélations, tendances — sont indiscernables de celles d'un portefeuille réel. Il s'agit d'une brique essentielle pour l'analyse quantitative en actuariat, permettant de surmonter les contraintes d'accès aux données. Le développement d'un tel outil s'est imposé comme une nécessité pour plusieurs raisons stratégiques et analytiques.

3.1.2 Besoins métiers : simulation de nouveaux produits et analyse concurrentielle

Pour un acteur du secteur de l'assurance, qu'il s'agisse d'un assureur ou d'un cabinet de conseil, la capacité à modéliser et à anticiper les dynamiques de marché est un avantage concurrentiel décisif. Le générateur de portefeuilles de passifs répond directement à ce besoin en fournissant un support quantitatif pour la prise de décision stratégique. Il permet par exemple à un cabinet de conseil de tester ses modèles sans dépendre des données clients, et à un assureur d'explorer des scénarios prospectifs ou d'évaluer l'impact de nouvelles offres. Son utilité se manifeste dans trois domaines clés pour les assureurs : le lancement de nouveaux produits, l'orientation du *business mix* et l'analyse concurrentielle.

Premièrement, le lancement d'un nouveau produit d'assurance-vie représente un investissement et un risque significatifs. Avant toute commercialisation, il est impératif d'en évaluer rigoureusement les impacts sur le profil de risque et la rentabilité de l'entreprise. Le générateur offre un véritable laboratoire virtuel pour effectuer ces tests. En simulant l'intégration de milliers de polices conformes aux caractéristiques du nouveau produit (garanties, frais, options), il permet de projeter leur comportement dans le temps.

Il devient alors possible d'analyser leur effet sur les indicateurs prudentiels de Solvabilité II, tels que le *Best Estimate* (BE) et le *Solvency Capital Requirement* (SCR), mais aussi d'évaluer leur sensibilité à divers chocs de marché (hausse des taux, krach boursier) ou de comportement (vagues de rachats). Cet outil permet ainsi de tester, d'ajuster et d'optimiser les caractéristiques d'un produit pour atteindre le couple rendement/risque désiré avant même sa mise sur le marché.

Deuxièmement, le générateur est un outil précieux pour piloter la stratégie à long terme de l'entreprise. La direction peut être amenée à vouloir faire évoluer son *business mix*, c'est-à-dire la répartition de son portefeuille entre différents types de produits (fonds en euros, unités de compte, prévoyance...). Par exemple, dans un contexte de taux bas persistants, un assureur pourrait vouloir accélérer sa transition vers les produits en unités de compte. Le générateur permet de quantifier les implications d'une telle stratégie. En simulant des portefeuilles futurs correspondant à ces nouvelles orientations commerciales, la direction peut visualiser les conséquences sur le bilan, la rentabilité prévisionnelle, mais aussi sur la consommation de capital et l'exposition aux risques. Ces simulations éclairent les décisions stratégiques et s'intègrent naturellement dans des exercices prospectifs comme l'ORSA (*Own Risk and Solvency Assessment*).

Enfin, la capacité à se positionner par rapport à ses concurrents est fondamentale. Faute d'accès aux portefeuilles détaillés des autres acteurs, un assureur doit s'appuyer sur des reconstitutions. En se basant sur des données publiques (rapports annuels ou publications réglementaires comme les SFCR) ou des statistiques sectorielles, le générateur peut permettre la création d'un portefeuille représentatif du marché, ou simuler le portefeuille probable d'un concurrent spécifique. Ces portefeuilles synthétiques deviennent alors une base solide pour des analyses comparatives (*benchmarking*). Ils permettent non seulement d'évaluer la performance relative, mais aussi de comparer les profils de risque, d'anticiper les stratégies concurrentes et d'identifier les meilleures pratiques du marché. Il convient toutefois de souligner les limites d'une telle démarche. Une analyse ALM complète et réaliste d'un concurrent ne peut se contenter de la seule modélisation du passif. Elle exigerait également de simuler son portefeuille d'actifs et de disposer d'informations précises sur ses ressources financières et ses fonds propres. Or, ces données, qui relèvent du secret des affaires, sont rarement publiques. Par conséquent, l'analyse comparative reste nécessairement partielle, se concentrant sur les caractéristiques intrinsèques du portefeuille de passifs reconstitué.

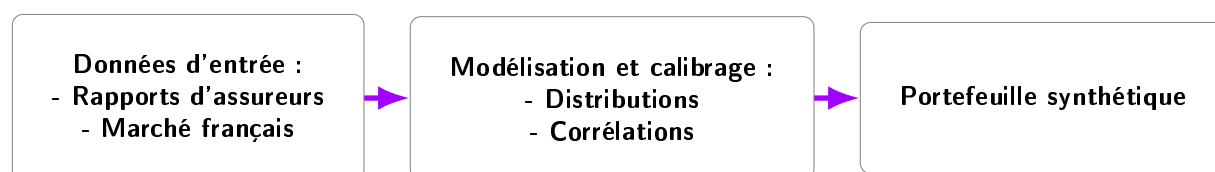


FIGURE 3.1 – Schéma de la méthodologie de génération d'un portefeuille de passifs synthétique.

3.1.3 Défis de la modélisation : réalisme, volumétrie et flexibilité

La conception et la mise en œuvre d'un générateur de portefeuilles de passifs efficace soulèvent trois défis majeurs et interdépendants :

- **Le réalisme des données générées :** Il s'agit du défi le plus fondamental. L'objectif n'est pas de produire des données aléatoires, mais de créer un portefeuille synthétique dont les propriétés statistiques sont indiscernables de celles d'un portefeuille réel. Cela implique non seulement de reproduire fidèlement les distributions de chaque caractéristique individuelle (âge, montant, etc.), mais aussi, et surtout, de capturer les corrélations complexes qui les lient. Par exemple, l'âge d'un assuré est souvent corrélé au type de produit souscrit et au montant de sa provision mathématique. Ignorer ces dépendances conduirait à un portefeuille incohérent, dont le comportement sous différents scénarios de risque serait erroné, invalidant ainsi les analyses prudentielles ou stratégiques qui en découlent.
- **La gestion de la volumétrie :** Les portefeuilles d'assurance-vie des grands acteurs du marché se comptent en centaines de milliers, voire en millions de contrats. Le générateur doit être capable de produire des ensembles de données de cette ampleur de manière performante, c'est-à-dire dans un temps de calcul raisonnable et sans consommer une quantité prohibitive de ressources mémoire. Cette contrainte de performance est d'autant plus forte que la gestion des corrélations, nécessaire au réalisme du portefeuille, ajoute une complexité de calcul significative. Il faut donc trouver un équilibre entre la sophistication statistique et la performance, ce qui a des implications directes sur les choix technologiques et algorithmiques.
- **La flexibilité de l'outil :** Un générateur ne serait que d'une utilité limitée s'il ne produisait qu'un seul type de portefeuille statique. Pour répondre aux besoins métiers variés, l'outil doit être hautement paramétrable. L'utilisateur doit pouvoir ajuster finement les caractéristiques du portefeuille à générer : définir les spécificités d'un nouveau produit, modifier les distributions statistiques pour simuler un segment de marché différent, ou encore changer les lois de comportement (rachat, mortalité) pour tester de nouvelles hypothèses. Cette flexibilité est la clé qui transforme le générateur en un véritable laboratoire d'expérimentation pour les actuaires et les stratèges.

Relever ces trois défis a nécessité de concevoir une architecture alliant rigueur statistique et performance calculatoire. La section suivante présente en détail la méthodologie de modélisation probabiliste qui a été développée pour construire un portefeuille à la fois réaliste, volumineux et flexible.

3.2 Méthodologie de Génération et Modélisation Statistique

La méthodologie de génération du portefeuille de passifs synthétique repose sur une approche probabiliste. L'objectif est de construire un ensemble de contrats d'assurance dont les propriétés statistiques sont entièrement maîtrisées. Pour ce faire, chaque caractéristique d'un contrat (âge de l'assuré, montant de la provision, etc.) est modélisée comme une variable aléatoire, tirée d'une loi de probabilité préalablement calibrée sur des données de marché quand elles sont disponibles. Sinon, des hypothèses ont été formulées pour recréer ces variables de la manière la plus réaliste possible.

3.2.1 Approche stochastique par lois de probabilité

Le cœur du générateur de passifs repose sur la modélisation de chaque attribut d'un contrat d'assurance vie par une loi de probabilité spécifique. Ce processus permet de créer une population de polices hétérogène et réaliste, dont la diversité est essentielle pour une simulation ALM pertinente. La génération s'effectue en deux temps : d'abord, la définition de "produits types" ou archétypes, puis la génération des contrats individuels à partir de ces archétypes. *Note pour moi-même : Le concept d'archétypes est une bonne idée, mais je dois le rendre plus concret. Je devrais définir précisément ce que c'est, combien j'en ai créé et sur quelle base (ex : "fonds en euros sécuritaire", "mixte dynamique", "pur UC"). Un tableau récapitulatif des archétypes et de leurs caractéristiques serait une excellente addition pour clarifier ce point.*

Les caractéristiques de ces profils (par exemple, les fourchettes de Taux Minimum Garanti (TMG), les montants moyens de provision mathématique) sont elles-mêmes générées aléatoirement à l'aide de lois uniformes, discrètes ou continues. Cette première étape permet de fixer les méta-paramètres (tels que la moyenne μ , l'écart-type σ , ou les bornes $[min, max]$) qui régiront la génération fine des contrats. *Note pour moi-même : Attention, ce passage est confus. Est-ce que je tire les paramètres des archétypes aléatoirement ? Si c'est le cas, je dois justifier cette couche de stochasticité. Sinon, si les paramètres sont fixés par des hypothèses expertes, je dois le dire clairement. Je devrais reformuler en précisant que "Les paramètres définissant chaque archétype (fourchettes de TMG, etc.) sont déterminés en amont..." pour plus de clarté.*

Chaque contrat individuel est ensuite instancié en tirant ses variables selon des lois de probabilité spécifiques, reflétant la nature de la caractéristique modélisée :

3.2.1.1 Modélisation de la distribution des âges

Une modélisation précise de la distribution de l'âge est essentielle lorsque l'on souhaite modéliser un portefeuille de passifs. En effet, c'est un facteur clé dans l'évaluation de risques tels que le décès. La démarche de modélisation s'est déroulée en quatre étapes principales : la préparation des données, l'ajustement de modèles, la comparaison statistique et la sélection du meilleur modèle.

La première étape a consisté à construire une distribution de probabilité réaliste à partir de deux sources

de données :

1. **La pyramide des âges de la population française** pour l'année 2024 [5], fournissant la structure démographique de base entre 0 et 100 ans.

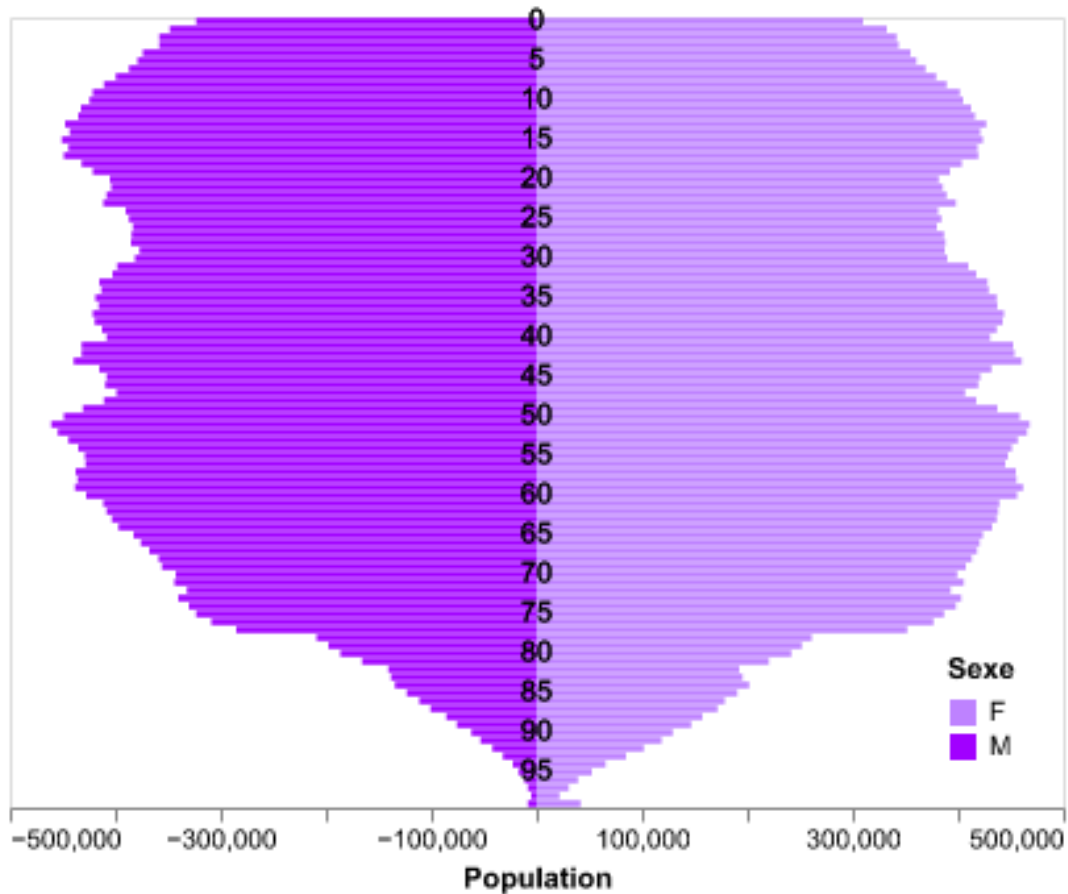


FIGURE 3.2 – Pyramide des âges de la population française en 2024 [5].

2. **Une étude statistique de l'INSEE** sur la détention d'assurance-vie par tranche d'âge en France [4]

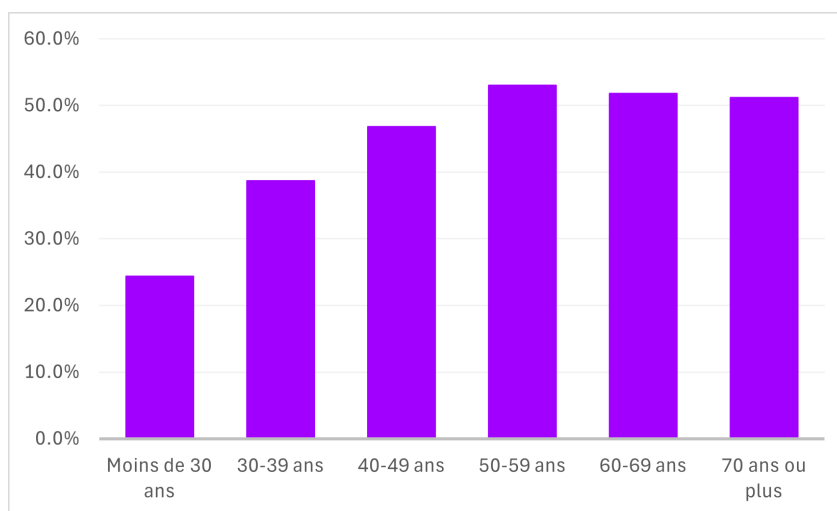


FIGURE 3.3 – Proportion de détention d'assurance-vie par tranche d'âge en France [4].

Afin d'obtenir de passer de la figure 3.3 à une proportion de personnes possédant une assurance vie par âge, il a été décidé de faire plusieurs interpolations entre les points de données fournis. Il a d'abord été décidé de définir des points précis pour chaque tranche d'âge, en prenant la moyenne de chaque tranche. Par exemple, pour la tranche 18-30 ans, le point retenu est 24 ans. Pour la dernière tranche "70 ans et plus", le point retenu est 82.5 ans. Ensuite une interpolation PCHIP a été réalisée entre les points de 24 ans à 64.5 ans. Pour atteindre le point suivant, 82.5 ans, une interpolation linéaire a été effectuée afin d'obtenir une décroissance réaliste. Puis après 82.5 ans, il a été décidé de garder la proportion constant car on prend en hypothèse que personne ne souscrit ou ne part de l'assurance vie après 82.5 ans. Ces interpolations permettent d'obtenir une estimation continue du taux de détention pour chaque âge entre 18 et 100 ans, qui est l'âge maximum retenu dans le cadre de la génération de portefeuilles dans le cadre de cette étude.

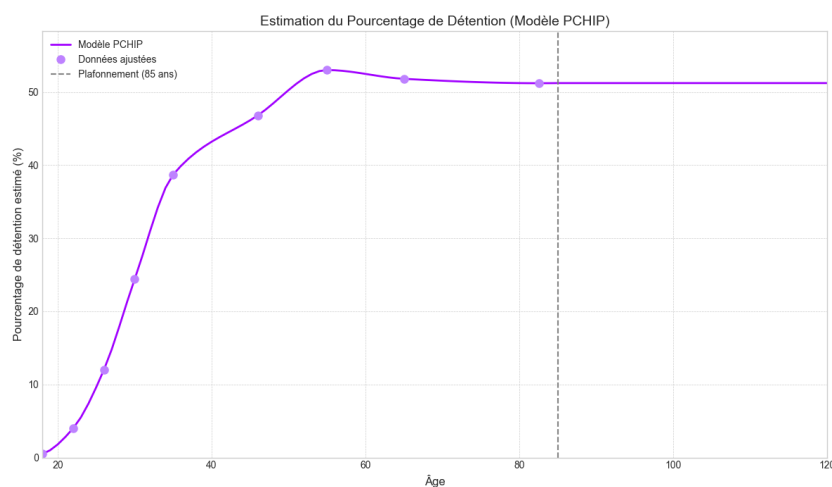


FIGURE 3.4 – Proportion de détention d'assurance-vie par âge en France par diverses méthodes d'interpolation.

En multipliant la population de chaque âge par le taux de détention estimé, il a été possible d'obtenir une estimation du nombre d'assurés pour chaque âge et chaque sexe. Après normalisation, il est désormais possible de construire une loi de probabilité empirique, illustrée par la figure 3.5.

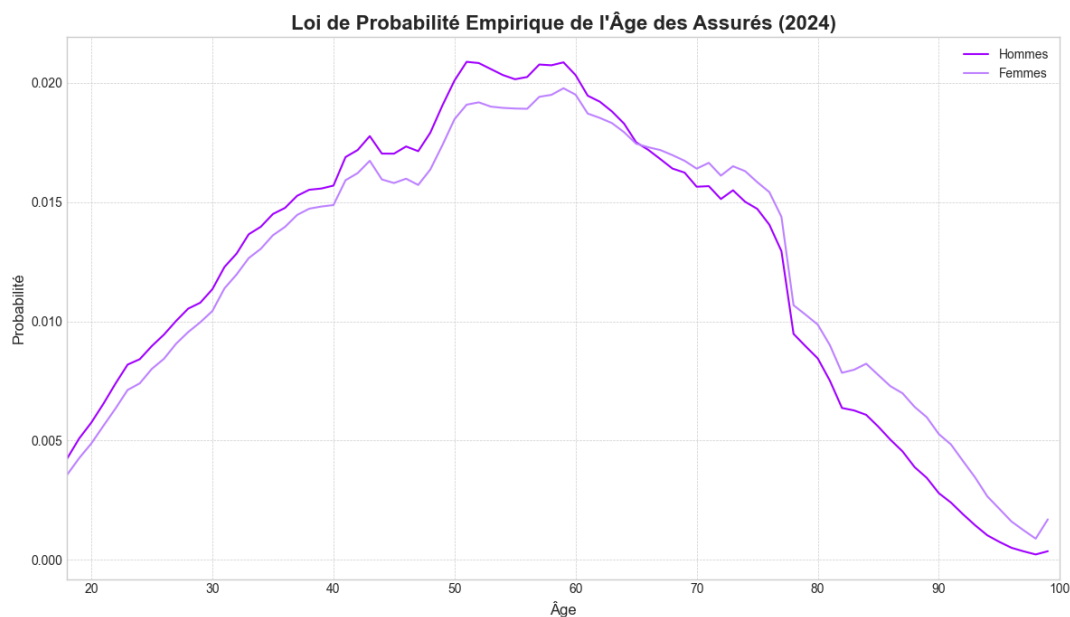


FIGURE 3.5 – Distribution empirique de l'âge des assurés (Hommes et Femmes).

Deux lois de probabilité continues ont été sélectionnées comme candidates pour modéliser la distribution empirique : la **loi Gamma** et la **loi Beta**. Les paramètres de ces deux lois ont été estimés par la méthode du maximum de vraisemblance sur un échantillon de 200 000 individus tirés de la loi empirique. Pour déterminer la loi la plus adéquate, des critères visuels et statistiques (Test de Kolmogorov-Smirnov, AIC, BIC) ont été utilisés. Les figures 3.6 et 3.7 montrent l'ajustement des deux lois. Visuellement, la loi Gamma épouse plus fidèlement la forme de la distribution observée.

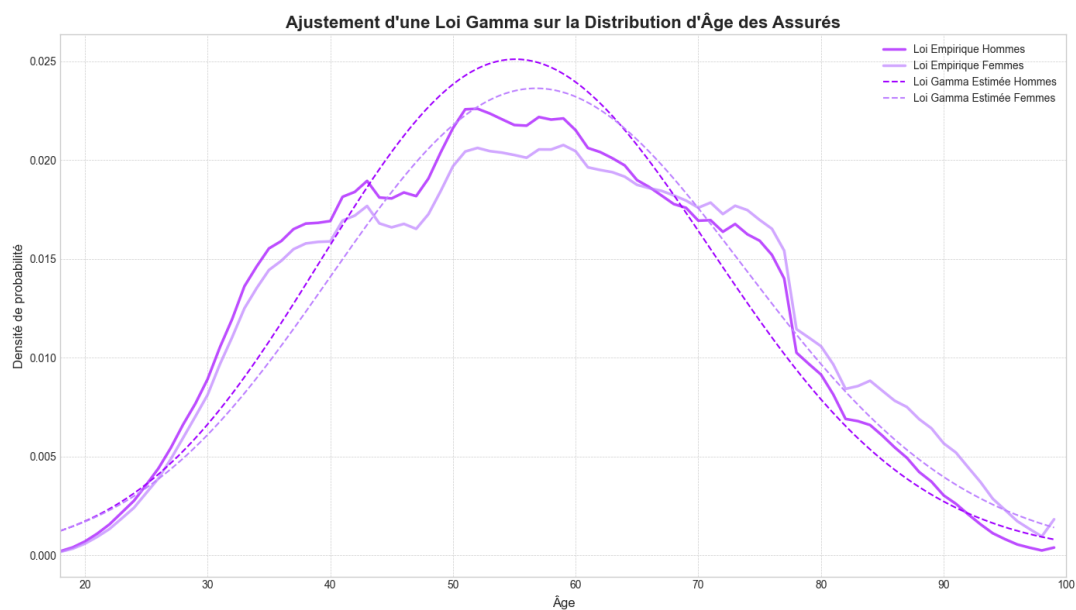


FIGURE 3.6 – Ajustement de la loi Gamma sur la distribution empirique.

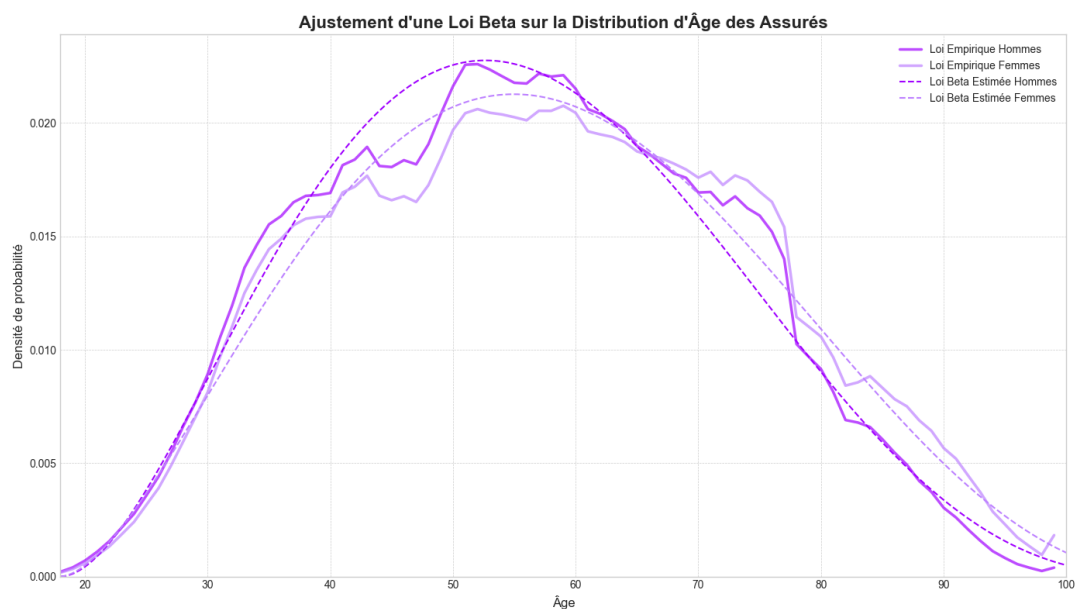


FIGURE 3.7 – Ajustement de la loi Beta sur la distribution empirique.

Le tableau 3.1 confirme cette observation. La loi Beta présente une statistique K-S deux fois plus faible et des scores AIC et BIC significativement inférieurs, indiquant un meilleur ajustement.

Métrique	Loi Gamma	Loi Beta	Meilleur Modèle
Statistique K-S (D)	0.0396	0.0271	Beta
AIC (Akaike)	1 674 497	1 668 357	Beta
BIC (Bayésien)	1 674 527	1 668 378	Beta

TABLE 3.1 – Tableau comparatif des métriques d'ajustement (population masculine).

D'après ces résultats, il est donc recommandé d'utiliser la **loi Beta**, avec les paramètres estimés lors de l'analyse, pour toute application nécessitant une génération ou une représentation de l'âge des assurés.

3.2.1.2 Modélisation de l'âge à la souscription

Après avoir modélisé l'âge des assurés, l'étape suivante consiste à déterminer l'âge auquel ils ont souscrit leur contrat. Cette variable est cruciale car elle nous permet de déterminer l'ancienneté du contrat, un facteur déterminant dans les outils de modélisations car il influence les taux de rachats.

La distribution de l'âge à la souscription a été dérivée des mêmes données de taux de détention par âge utilisées précédemment. En effet, la probabilité de souscrire à un âge donné peut être approximée par la variation (la dérivée discrète) du taux de détention. Une augmentation rapide du taux de détention entre 25 et 35 ans, par exemple, implique une forte probabilité de souscription dans cette tranche d'âge.

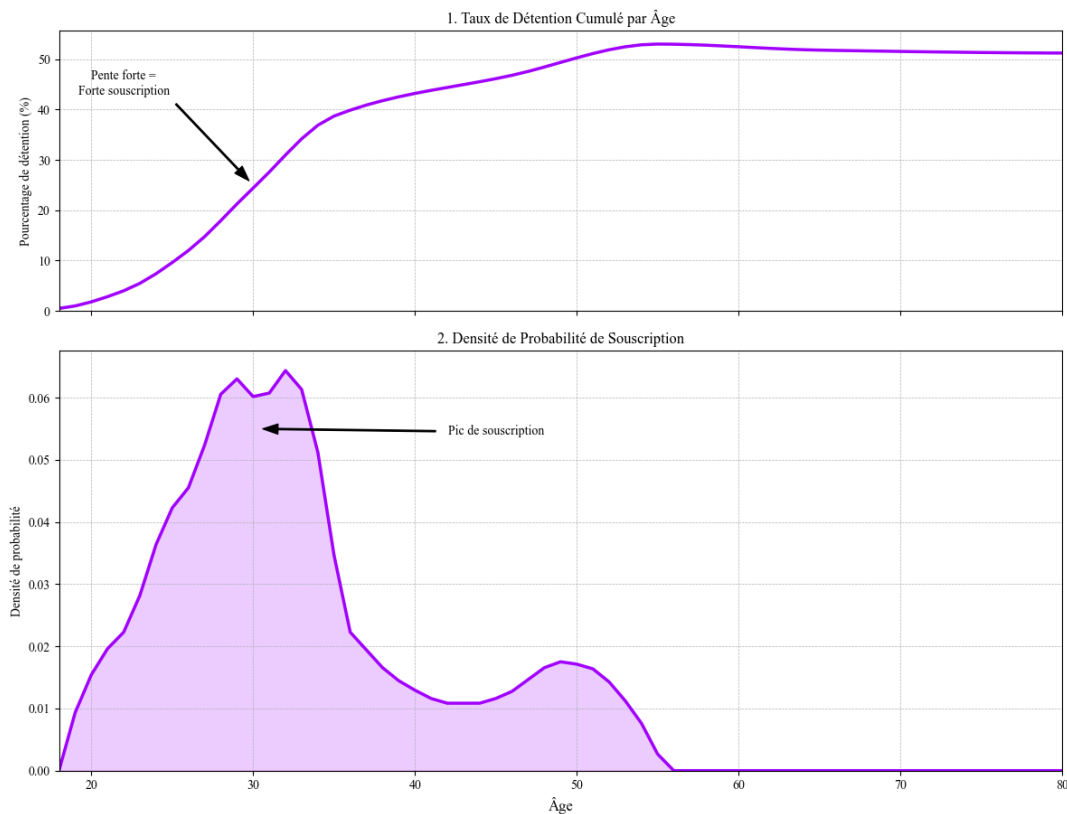


FIGURE 3.8 – Passage de la distribution empirique de l'âge à la souscription à la densité de probabilité du taux de souscription.

Plusieurs lois de probabilité ont été testées pour modéliser cette distribution empirique de l'âge à la souscription : Gamma, Beta, Weibull, et un modèle de Mélange Gaussien (GMM). Le tableau 3.2 présente les résultats.

Distribution	AIC	BIC	K-S (D)
GMM (n=2)	1 125 603	1 125 654	0.0485
Beta	1 385 454	1 385 495	0.0845
Gamma	1 400 178	1 400 208	0.1110
Weibull	1 437 176	1 437 207	0.1544

TABLE 3.2 – Tableau comparatif des métriques d'ajustement pour l'âge à la souscription.

Le **Mélange Gaussien à deux composantes (GMM)** s'est avéré être le modèle le plus performant, avec des scores AIC et BIC nettement inférieurs. D'après ce résultat, nous pouvons remarquer qu'il y a alors 2 pics de souscription. Cette bimodalité, illustrée en figure 3.9, est une caractéristique de marché que les lois plus simples ne peuvent capturer.

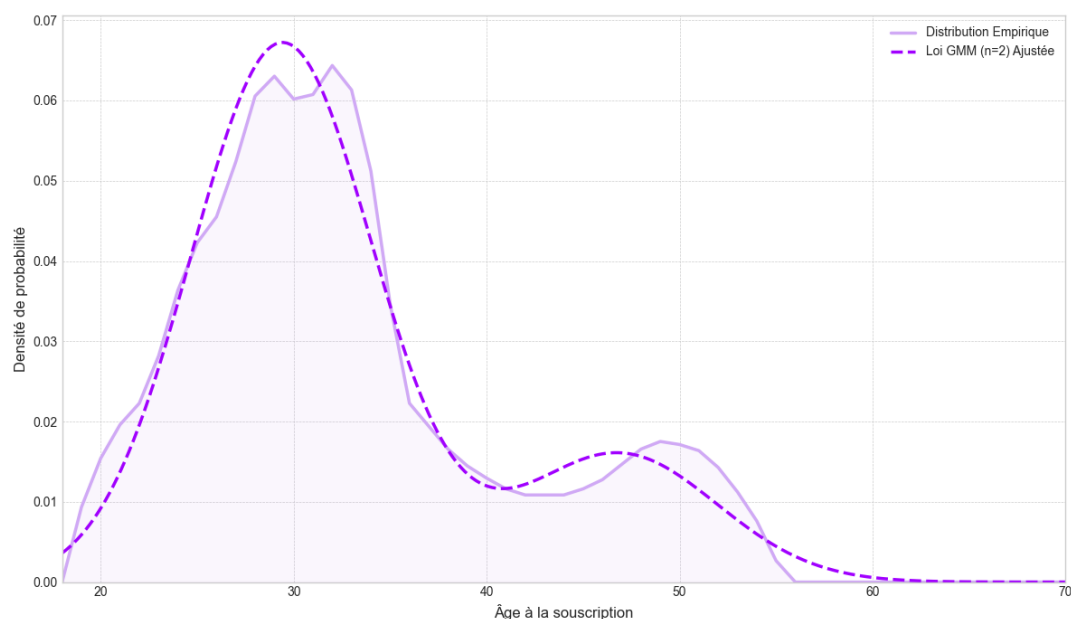


FIGURE 3.9 – Ajustement d'une GMM sur la distribution de l'âge à la souscription.

3.2.1.3 Modélisation de la Provision Mathématique (PM)

Pour calculer la Provision Mathématique à partir des données marché, une méthode plus complexe a été mise en place. Une approche directe consistant à ajuster une loi sur des données de PM n'est pas possible car des données publiques sur ce sujet n'existent pas. Une méthodologie de modélisation conditionnelle a donc été mise en place.

L'hypothèse fondamentale est que la PM d'un individu est principalement fonction de son **patrimoine**, qui lui-même est fortement corrélé à son **âge**. La modélisation s'est donc déroulée en plusieurs étapes.

La première étape a consisté à modéliser la distribution du patrimoine brut en fonction de l'âge, en s'appuyant sur les données de l'INSEE [7]. Plutôt que de calibrer une seule loi pour toute la population, une **loi Lognormale** a été ajustée pour chaque tranche d'âge. Les paramètres de cette loi, μ et σ , ont été estimés par la méthode de **correspondance des quantiles**, en s'assurant que la médiane et les déciles de la loi ajustée correspondent à ceux observés dans les données. Le tableau 3.3 synthétise les paramètres obtenus.

Tranche d'âge	μ	σ
Moins de 30 ans	9.9233	1.7713
30 à 39 ans	11.6750	1.8374
40 à 49 ans	12.1756	2.1221
50 à 59 ans	12.3216	2.0551
60 à 69 ans	12.3579	2.0772
70 ans ou plus	12.2620	1.8199

TABLE 3.3 – Paramètres de la loi Lognormale du patrimoine brut, calibrés par tranche d'âge.

Sur la base des calibrations précédentes, une population synthétique de 500 000 individus a été générée. Pour chaque individu, un âge a été tiré selon la loi Beta, puis un patrimoine a été tiré selon la loi Lognormale conditionnelle correspondant à son âge. On obtient ainsi un échantillon de paires (Âge, Patrimoine) respectant la corrélation observée dans la réalité.

La deuxième étape consiste à estimer la part du patrimoine de chaque individu allouée à l'assurance-vie. Pour ce faire, les données de l'INSEE sur la composition du patrimoine par décile [7] sont utilisées. La figure 3.10 illustre la part du patrimoine financier en fonction du patrimoine brut moyen. Cette relation n'est pas linéaire. Pour les patrimoines les plus modestes, qui n'ont pas encore la capacité d'investir dans l'immobilier, la part des actifs financiers est relativement élevée. À mesure que le patrimoine augmente, une part significative est allouée à l'immobilier, ce qui diminue mécaniquement la part relative du patrimoine financier. Ce phénomène s'inverse cependant pour les 10% des patrimoines les plus élevés, où la diversification vers des actifs financiers redevient prépondérante, entraînant une remontée de leur part dans le patrimoine total.

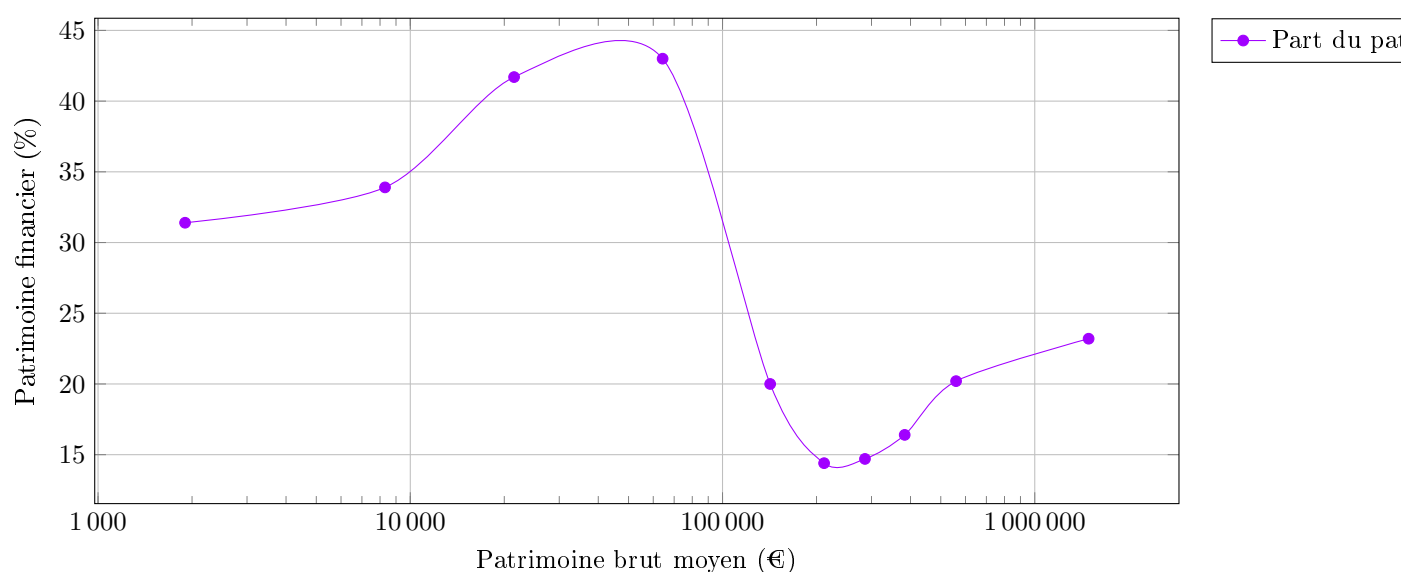


FIGURE 3.10 – Part du patrimoine financier en fonction du patrimoine brut moyen, par décile [7].

La troisième étape consiste à lier le patrimoine financier à la Provision Mathématique. Il est en effet plus réaliste de considérer que l'épargne en assurance-vie (ou la provision mathématique du point de vue de l'assureur) constitue une part du patrimoine financier plutôt que du patrimoine brut. Une hypothèse centrale est donc formulée : la PM d'un individu est estimée comme une fraction de son patrimoine financier. Sur la base de données INSEE, cette fraction est fixée à **40%**. Ainsi, pour chaque individu de la population simulée, la PM est calculée comme suit :

$$PM = \text{Patrimoine Brut} \times \text{Part du Patrimoine Financier} \times 40\%$$

Cette règle permet de transformer la distribution de patrimoine en une distribution de PM, en tenant

compte des non-linéarités observées dans la composition du patrimoine.

Après avoir appliqué ce processus à toute la population simulée, nous obtenons un échantillon réaliste de Provisions Mathématiques. Une dernière calibration a montré que la distribution de ces PM pouvait être modélisée de manière très satisfaisante par une **loi Lognormale**, comme l'indique le tableau 3.4.

Distribution	AIC	BIC	K-S (D)
Lognorm	7 142 088	7 142 120	0.017
Gamma	7 254 743	7 254 774	0.145

TABLE 3.4 – Pour le critère de Kolmogorov-Smirnov, la loi Lognormale est largement meilleure que la loi Gamma pour estimer cette distribution.

3.2.2 Synthèse et Génération du Portefeuille Final

Avec toutes les distributions marginales et conditionnelles calibrées, l'étape finale consiste à les assembler pour générer le portefeuille complet. Une approche par copule classique, où les trois variables (Âge, Âge Souscription, PM) seraient mises dans une copule 3D, a été écartée. Une telle approche simplifierait à l'excès les dépendances, en les réduisant à de simples coefficients de corrélation, et ne saurait capturer la richesse des relations non-linéaires (notamment l'évolution du patrimoine avec l'âge).

Un **modèle de génération hybride** a donc été privilégié :

1. **Génération des âges par copule** : Une copule Gaussienne 2D est utilisée pour générer des paires corrélées (Âge de l'assuré, Âge de souscription), en utilisant les lois marginales calibrées (Gamma et GMM). Une contrainte assure que l'âge de souscription est toujours inférieur à l'âge de l'assuré.
2. **Génération conditionnelle de la PM** : Pour chaque couple d'âges généré, la cascade de simulation décrite dans la section précédente est appliquée :
 - On génère un patrimoine à partir de la loi Lognormale correspondant à l'âge de l'assuré.
 - On simule la détention d'un contrat en fonction de ce patrimoine.
 - Si un contrat est détenu, on calcule la PM en se basant sur la part de patrimoine financier et la règle des 40%.

Ce processus itératif est répété jusqu'à obtenir le nombre de contrats désiré dans le portefeuille final. Le résultat est un fichier de données synthétique contenant trois variables dont les distributions et les interdépendances sont le reflet fidèle des analyses statistiques menées en amont. Cet outil fournit ainsi une base de données robuste et réaliste pour toutes les analyses actuarielles ultérieures.

- **Nombre de polices par agrégat** : Modélisé par une loi de Poisson $\mathcal{P}(\lambda)$, tronquée pour assurer une valeur minimale de 1. *Note pour moi-même : Je dois justifier le choix de la loi de Poisson. C'est un choix classique, mais je dois expliquer pourquoi je l'ai fait (ex : grand nombre de clients potentiels, faible probabilité de souscription). Je dois aussi préciser comment le paramètre lambda*

est calibré. Pour la "troncature à 1", je dois être plus précis mathématiquement. Est-ce que j'utilise une loi de Poisson tronquée en zéro (Zero-Truncated Poisson) ? Si oui, je devrais donner la fonction de masse correspondante.

- **Sexe de l'assuré :** Modélisé par une loi de Bernoulli $\mathcal{B}(p)$, où p représente la proportion d'hommes définie dans les spécifications du produit. *Note pour moi-même : C'est un choix standard. Je dois juste préciser d'où vient le paramètre p . Est-ce que je l'ai tiré des données de marché, des statistiques nationales (INSEE), ou est-ce une hypothèse propre au produit cible ?*
- **Montants des provisions (fonds euros et UC) :** Tirés selon une loi Normale $\mathcal{N}(\mu, \sigma^2)$, dont les paramètres sont issus des spécifications. Les valeurs générées sont ensuite tronquées pour rester dans des bornes prédéfinies, afin d'éviter les valeurs aberrantes. *Note pour moi-même : Le choix de la loi Normale pour des montants est discutable. Une loi Lognormale ou Gamma serait peut-être plus adaptée. Je dois justifier ce choix. Ai-je fait des tests d'ajustement (Kolmogorov-Smirnov, etc.) sur des données réelles ? Je dois aussi détailler la troncature : quelles bornes $[\min, \max]$ et comment ont-elles été choisies ? Je dois vérifier si j'ai bien tenu compte de l'impact de la troncature sur la moyenne et la variance pour que les moments finaux correspondent aux cibles.*
- **Allocation d'actifs cible :** Les poids cibles pour chaque classe d'actifs (actions, immobilier, etc.) sont également générés suivant des lois Normales. Un mécanisme de normalisation est ensuite appliqué pour garantir que la somme des allocations $\sum_i \omega_i$ pour un contrat donné est égale à 1. *Note pour moi-même : Même remarque que pour les provisions. La normalisation post-tirage modifie les lois marginales. Une approche plus rigoureuse serait d'utiliser une loi de Dirichlet. Je devrais mentionner cette alternative et justifier pourquoi je ne l'ai pas utilisée (par exemple, pour des raisons de simplicité).*
- **Date d'effet du contrat :** Générée par une loi Uniforme continue sur un intervalle de temps $[T_{\text{début}}, T_{\text{fin}}]$. *Note pour moi-même : C'est un choix simple. Je dois le justifier brièvement. Est-ce que l'hypothèse d'une souscription uniforme dans le temps est réaliste pour mon cas ? Je pourrais mentionner que d'autres lois comme Beta ou Triangulaire pourraient mieux modéliser des pics de souscription.*

3.2.3 Calibration des distributions marginales à partir des données de marché

L'approche stochastique adoptée n'est pas purement aléatoire; elle est rigoureusement calibrée pour refléter une réalité de marché ou un portefeuille cible. La calibration des distributions marginales de chaque variable est assurée principalement via un ensemble de spécifications qui définissent les paramètres des lois de probabilité. Ces spécifications peuvent être obtenues de deux manières complémentaires :

1. **Calibration par hypothèses expertes :** Les paramètres des lois (moyennes, écarts-types, bornes) sont définis sur la base d'hypothèses actuarielles et économiques. Cette méthode est utilisée pour explorer des scénarios de marché spécifiques ou pour concevoir des produits cibles. *Note pour moi-*

même : Ce paragraphe est trop vague, il faut que je le rende plus concret avec des exemples précis.

2. **Calibration directe sur données réelles :** Les spécifications peuvent être issues d'une analyse statistique d'un portefeuille existant. Les paramètres des lois sont alors les estimateurs directs des moments empiriques (moyenne, variance) observés sur les données, garantissant que le portefeuille synthétique réplique fidèlement la structure du portefeuille réel. *Note pour moi-même : C'est le point le plus faible de la section, je dois absolument le renforcer. Je dois décrire précisément la source de données utilisée (portefeuille réel anonymisé ? taille ? période ?). Je dois expliquer la méthode d'estimation (méthode des moments ou maximum de vraisemblance ? Pourquoi ce choix ?). Surtout, je dois ajouter la validation statistique : quels tests d'ajustement (goodness-of-fit) ai-je utilisés (Chi-deux, K-S) et quels sont les résultats (p-valeurs) ? Des graphiques comparant l'histogramme des données réelles et la densité de la loi calibrée sont indispensables ici.*

Un exemple particulièrement abouti de cette calibration est la modélisation de la distribution des âges des assurés. Le processus charge une distribution d'âge de référence, représentative de la population cible. Cette distribution est ensuite transformée en une série de lois de probabilité conditionnelles : pour chaque âge maximum possible à la souscription, n , une distribution de probabilité discrète, tronquée à n , est calculée. *Note pour moi-même : C'est un bon exemple, mais je dois être plus précis. Quelle est la "distribution d'âge de référence" ? Table de mortalité réglementaire, données INSEE ? Je dois citer la source exacte. Je devrais aussi décrire le processus mathématiquement. Si $f(x)$ est la densité de référence, la densité conditionnelle est $f_{X|X \leq n}(x) = \frac{f(x)}{\int_0^n f(t) dt}$. Montrer cette formule prouverait ma maîtrise du sujet.* Lors de la génération, l'âge d'un assuré pour un contrat donné est tiré au hasard en utilisant la distribution de probabilité conditionnelle qui respecte l'âge maximum autorisé pour ce produit. Cette méthode garantit que la structure démographique du portefeuille généré est cohérente avec celle de la population de référence.

3.2.4 Perspective : modélisation des dépendances par la théorie des copules

Note pour moi-même : Cette partie devrait plutôt aller dans une conclusion pour donner une ouverture à mon sujet de mémoire Une limite inhérente au modèle actuel est l'hypothèse d'indépendance entre les différentes caractéristiques des contrats. Par exemple, le montant de la provision mathématique est généré indépendamment de l'âge de l'assuré, ce qui est une simplification forte de la réalité. Pour dépasser cette limite, une évolution naturelle du modèle consisterait à intégrer la théorie des copules pour modéliser la structure de dépendance entre les variables aléatoires.

La force de l'approche par copules réside dans sa capacité, via le théorème de Sklar, à séparer la modélisation des distributions marginales de celle de leur structure de dépendance. Il serait ainsi possible de conserver la calibration fine de chaque marge, décrite précédemment, tout en introduisant une structure de corrélation réaliste. La mise en œuvre suivrait les étapes suivantes :

1. **Définition de la structure de dépendance :** Spécification d'une matrice de corrélation (par exemple, de Spearman ou de Kendall) entre les variables d'intérêt, comme l'âge, l'ancienneté du contrat et le montant des provisions. *Note pour moi-même : C'est une bonne perspective. Pour la renforcer, je devrais expliquer comment je calibrerais cette matrice de corrélation (estimation à partir des rangs sur les données réelles ?). Je pourrais aussi préciser que le choix entre Spearman*

et Kendall n'est pas anodin.

2. **Choix d'une famille de copules :** Sélection d'une copule (Gaussienne, de Student, ou archimédienne comme celles de Clayton ou Gumbel) en fonction de la nature des dépendances à modéliser, notamment la symétrie ou la dépendance aux extrêmes (risques de queue). *Note pour moi-même : Parfait. Pour aller plus loin, je pourrais donner un exemple concret : "Par exemple, pour modéliser le fait que les assurés les plus âgés ont tendance à avoir les plus grosses provisions (dépendance dans la queue supérieure), une copule de Gumbel serait plus appropriée qu'une copule Gaussienne." Cela montrerait une compréhension plus fine.*
3. **Modification du processus de génération :** Le tirage des variables ne se ferait plus de manière indépendante. Il faudrait d'abord générer des vecteurs de variables uniformes corrélées à partir de la copule choisie. Ensuite, par la méthode de la transformée inverse, appliquer la fonction de répartition inverse de chaque loi marginale calibrée ($F_X^{-1}(u)$) pour obtenir les réalisations des variables finales. *Note pour moi-même : Je dois réfléchir à la faisabilité d'appliquer réellement cette méthode dans le cadre du mémoire.*

Cette extension permettrait d'enrichir considérablement le réalisme du portefeuille synthétique, en capturant des phénomènes clés tels que la tendance des assurés plus âgés à détenir des contrats avec des provisions plus élevées, améliorant ainsi la pertinence des simulations ALM subséquentes.

3.3 Présentation de l'Outil et du Portefeuille de Référence Généré

3.3.1 Architecture de l'application en Python et choix technologiques

3.3.2 Description des paramètres d'entrée et des formats de sortie

3.3.3 Analyse descriptive du portefeuille de référence

Chapitre 4

Protocole d'Analyse : Sélection d'une Méthode d'Agrégation

Ce chapitre a pour objectif de définir et de mettre en œuvre un protocole d'analyse rigoureux visant à sélectionner la méthode d'agrégation la plus adaptée à notre étude. Le choix s'appuiera sur une évaluation comparative de plusieurs approches candidates, en considérant des critères clés tels que la fidélité des indicateurs prudentiels (BE et SCR), la performance en termes de temps de calcul et la complexité de mise en œuvre.

La démarche consistera d'abord à présenter les différentes méthodes d'agrégation envisagées. Ensuite, nous appliquerons chacune de ces méthodes à un portefeuille de référence afin de quantifier les écarts de résultats par rapport à une modélisation sans agrégation.

À l'issue de cette analyse comparative, la méthode jugée la plus performante sera retenue. Elle sera ensuite utilisée pour conduire les analyses de sensibilité détaillées dans le chapitre suivant.

4.1 Présentation des Méthodes d'Agrégation candidates

4.1.1 Approches par clustering (K-means, DBSCAN/HDBSCAN)

4.1.2 Autres approches (MP par âge, MP Mémoire Amine Ben Fadhel, etc.)

4.2 Définition du Protocole de Test Comparatif

4.2.1 Constitution des portefeuilles de test

4.2.2 Définition des critères de sélection : fidélité des indicateurs (BE/SCR), performance et temps de calcul

4.3 Analyse Comparative et Choix de la Méthode Optimale

4.3.1 Synthèse des performances pour chaque méthode candidate

4.3.2 Justification du choix de la méthode retenue pour l'analyse de sensibilité

Chapitre 5

Analyse de Sensibilité des Indicateurs S2 aux Caractéristiques du Passif

Le but de ce chapitre est de tester la robustesse de la méthode d'agrégation sélectionnée au chapitre 4. Pour ce faire, des chocs (réglementaires + autres) seront appliqués sur des portefeuilles de test générés via le générateur présenté au chapitre 3. L'objectif est d'analyser l'impact de ces chocs sur les indicateurs S2 (BE, SCR) en comparant les résultats obtenus sur les portefeuilles granulaires (sans agrégation) et les portefeuilles agrégés (avec la méthode d'agrégation retenue). Cette analyse permettra de vérifier que la méthode d'agrégation choisie préserve la sensibilité des indicateurs aux variations des caractéristiques du passif, assurant ainsi la fiabilité des analyses de risque et de solvabilité dans un contexte d'optimisation computationnelle.

5.1 Définition des Scénarios de Sensibilité

5.1.1 Création des portefeuilles de test via le générateur

5.1.2 Description des chocs sur les variables clés (âge, montant de la PM, etc.)

5.1.3 Scénario d'intégration d'un nouveau produit dans le portefeuille

5.2 Analyse de l'Impact de l'Agrégation sur la Mesure des Chocs

- 5.2.1 Comparaison des indicateurs S2 sur portefeuilles choqués granulaires et agrégés
- 5.2.2 Analyse de la fidélité de la méthode d'agrégation à retranscrire la sensibilité

5.3 Interprétation des Résultats et Validation de l'Approche

- 5.3.1 Validation de la performance de la chaîne de modélisation (Générateur - Agrégation - Modèle ALM)
- 5.3.2 Enseignements sur la sensibilité des portefeuilles aux modifications de caractéristiques du passif

Chapitre 6

Conclusion

6.1 Résumé des résultats

6.1.1 Synthèse des principaux résultats obtenus

6.1.2 Impact des méthodes d'agrégation sur les portefeuilles de passifs

6.2 Perspectives d'amélioration

- 6.2.1 Axes d'amélioration pour les générateurs de portefeuilles de passifs
- 6.2.2 Évolutions possibles des méthodes d'agrégation et de modélisation ALM
- 6.2.3 Autres domaines d'application des générateurs de portefeuilles de passifs

6.3 Conclusion Générale et Perspectives

Ce mémoire, présenté ici dans sa version intermédiaire, dresse le bilan d'une première phase de recherche consacrée à la construction d'un cadre méthodologique et au développement des outils de modélisation. Les résultats quantitatifs finaux ne sont pas encore établis à ce stade, la priorité ayant été donnée à l'élaboration d'un socle technique robuste. Cette conclusion a donc pour double vocation de définir avec précision la feuille de route pour la finalisation des travaux et de présenter un bilan synthétique des compétences clés acquises chez Accenture, qui ont été déterminantes pour ce projet.

6.3.1 Perspectives et Finalisation des Travaux de Recherche

Le modèle ALM sous Python, désormais opérationnel, constitue la pierre angulaire de l'analyse à venir. Les prochaines étapes seront dédiées à son exploitation intensive afin de répondre à la problématique du mémoire. Les axes de travail prioritaires sont les suivants :

- **Analyse Comparative et Critique des Méthodes d'Agrégation** : Une investigation approfondie de plusieurs méthodes d'agrégation de portefeuilles sera menée. Au-delà d'une simple application mécanique, il s'agira de disséquer les fondements théoriques de chaque approche, leurs hypothèses et leur incidence sur la mesure de la diversification des risques. Une attention particulière sera portée à la **robustesse** des méthodes, c'est-à-dire leur capacité à produire des indicateurs agrégés stables et convergents, y compris sous des conditions de marché dégradées ou lors de variations des hypothèses de modélisation.
- **Enrichissement du Générateur de Portefeuilles de Passifs** : La crédibilité des résultats repose sur le réalisme des simulations. Le générateur de portefeuilles sera donc affiné pour intégrer des structures de dépendance plus sophistiquées entre les variables clés (âge des assurés, montant des provisions mathématiques, comportement de rachat, etc.). Cette évolution permettra de simuler des portefeuilles plus fidèles à la réalité d'un assureur et de tester la résilience de nos conclusions dans des scénarios adverses.
- **Conduite d'une Campagne d'Analyses de Sensibilité** : Une fois les méthodologies d'agrégation validées, une série de tests de sensibilité sera déployée. En appliquant les chocs réglementaires prévus par Solvabilité 2 (choc taux, actions, mortalité, rachat), nous vérifierons la stabilité des agrégations et quantifierons leur impact sur le Solvency Capital Requirement (SCR). L'objectif final est d'offrir une vision claire de la manière dont l'agrégation modifie et stabilise le profil de risque d'un portefeuille.

La feuille de route pour la finalisation du mémoire est clairement établie et s'inscrit dans la continuité logique des travaux déjà réalisés. Les parties non encore écrites ont un plan clair et détaillé qui montre la robustesse de l'approche. La maîtrise des outils et des concepts acquise jusqu'à présent me permet d'aborder cette dernière phase avec une grande confiance. Une planification rigoureuse et une exécution méthodique assureront l'achèvement des analyses et la rédaction finale dans les délais impartis.

6.3.2 Bilan de l'expérience professionnelle chez Accenture : un levier pour le mémoire

Mon année en alternance chez Accenture a été une expérience fondatrice, m'offrant un cadre idéal pour développer les compétences techniques indispensables à la réalisation de ce mémoire.

6.3.2.1 Développement et industrialisation d'un modèle ALM en Python

Au cœur de mes missions se trouvait le développement d'un modèle ALM (Asset-Liability Management) entièrement programmé en Python. Sous la supervision d'un Manager, j'ai pu assimiler les meilleures pratiques de l'ingénierie logicielle appliquée à l'actuariat. Cela inclut la gestion de projet via **Git**, un système de contrôle de version essentiel pour le travail collaboratif et le suivi rigoureux des modifications du code. J'ai également mis en place des batteries de tests unitaires, notamment via des interfaces avec Excel, pour garantir la fiabilité et la cohérence des résultats du modèle à chaque itération.

Ce socle technique m'a permis de mener à bien des projets complexes, comme la transformation du modèle ALM développé en un modèle de projection de passifs de retraite. Cette adaptation a exigé une refonte profonde de la logique de modélisation du passif, consolidant ainsi ma compréhension des mécanismes actuariels. J'ai par ailleurs conçu et développé un générateur de portefeuilles de passifs, un outil stratégique pour Accenture afin de réaliser des tests de performance et de résistance sur le modèle.

6.3.2.2 Développement d'un modèle de passifs sociaux sous Python

Dans le cadre de mon alternance chez Accenture, j'ai eu l'opportunité de recréer un logiciel calculant des passifs sociaux en Python. Les passifs sociaux sont des engagements financiers liés aux prestations sociales, tels que les retraites ou les indemnités de départ à la retraite que les entreprises doivent provisionner. Ce projet m'a permis de mettre en pratique mes compétences en programmation tout en approfondissant ma compréhension des enjeux actuels liés à la gestion des ressources humaines et à la prévoyance sociale.

Ce logiciel a été conçu pour être modulable et facilement intégrable dans les systèmes existants des clients d'Accenture, facilitant ainsi son adoption. J'ai également veillé à ce qu'il soit accompagné d'une documentation complète, permettant aux utilisateurs de comprendre rapidement son fonctionnement et de l'utiliser efficacement pour des futures missions sur ce sujet.

6.3.2.3 Maîtrise d'un écosystème technique moderne

Cette immersion professionnelle m'a permis de maîtriser un environnement de développement avancé. J'ai travaillé quotidiennement sous **Linux**, un système d'exploitation open-source prisé pour sa stabilité et sa sécurité, qui est le standard dans de nombreux environnements de calcul scientifique et de serveurs. J'ai également acquis une compétence sur **Docker**, une technologie de conteneurisation. Un conteneur est une sorte de "boîte" logicielle qui embarque une application et toutes ses dépendances, garantissant ainsi qu'elle s'exécute de manière identique et fiable, quel que soit l'ordinateur ou le serveur sur lequel elle est déployée.

6.3.2.4 Expertise en gestion de données et en calcul distribué (Cloud)

La manipulation de grands volumes de données est au cœur des problématiques actuarielles modernes. J'ai pu renforcer mes compétences en utilisant des bibliothèques Python de haute performance comme **Polars** pour le traitement de données et **Xlwings** pour automatiser les interactions entre Python et Excel. Mon apprentissage le plus significatif fut la prise en main de **Snowflake**, une plateforme de données hébergée dans le cloud. Snowflake permet non seulement de stocker et de requêter d'énormes jeux de données via le langage SQL, mais aussi d'exploiter sa puissance de calcul massive pour exécuter des scripts Python, comme le modèle ALM, directement sur la plateforme, optimisant ainsi drastiquement les temps de traitement et les besoins en ressources.

6.3.2.5 Collaboration, autonomie et transmission des savoirs

Au-delà des aspects purement techniques, j'ai appris à évoluer au sein d'une équipe projet agile, en participant à des réunions de suivi quotidiennes. J'ai également eu la responsabilité de former plusieurs nouveaux collaborateurs, à la fois sur le fonctionnement technique du modèle ALM et sur les bonnes pratiques de développement à adopter. Cette expérience de transmission a été extrêmement formatrice, renforçant mes capacités de communication et de vulgarisation de concepts complexes.



Annexes

Bibliographie

- [1] DORNAIKA Fadi, HAMAD Denis, CONSTANTIN Joseph, TRONG HOANG Vinh. *Advances in Data Clustering*. Lieu d'édition : Springer, 2024.
- [2] GOFFARD Pierre-Olivier, GUERRAULT Xavier. « Is it optimal to group policyholders by age, gender, and seniority for BEL computations based on model points? ». *European Actuariel Journal*, volume 5, 17 Avril 2015, p. 165-180.
- [3] BEN FADHEL Amine. « Accélération de l'évaluation de la solvabilité prospective d'un assureur épargne ». *Mémoire pour l'Institut des Actuaire*s , 2022.
- [4] INSEE. *Taux de détention des produits de patrimoine selon l'âge de la personne de référence du ménage en 2021*. (2021). Consulté le 18 Septembre 2025, sur https://www.insee.fr/fr/outil-interactif/5367857/details/30_RPC/34_PAT/34A_Figure1
- [5] INSEE. *Pyramide des âges de la population française en 2024*. (2024). Consulté le 1er Septembre 2025, sur <https://www.insee.fr/fr/outil-interactif/5014911/pyramide.htm>
- [6] FRANCE ASSUREURS. *L'assurance vie en 2023*. (20 septembre 2024). Consulté le 1er Septembre 2025, sur <https://www.franceassureurs.fr/nos-chiffres-cles/assurance-vie/etude-statistique-assurance-vie-2023/>
- [7] INSEE. *Enquête Histoire de vie et Patrimoine 2020-2021*. 2021.