



Master Économiste de l'entreprise - Université de Tours

2025–2026

3R : RECHERCHE, RÉALISATION, RÉSTITUTION

Gestion de portefeuille avec risques multiples

Appliquée à un portefeuille d'assurance automobile

Réalisé par

Nora ROUSSEAU, Lyna BOUDAMOUS, Alexis VINCENT et Emmanuel PAGUIEL

Dirigé par

M. Alain BOUSQUET

Décembre 2025

Table des matières

1	Fondements de la gestion de portefeuille	4
1.1	Introduction	4
1.2	Présentation de la théorie de Markowitz	5
1.2.1	Revue de Littérature	5
1.2.2	Origine et fondements de la théorie moderne du portefeuille	6
1.2.3	La Question Révolutionnaire de Markowitz	6
1.3	Concepts Fondamentaux	7
1.3.1	Risque et Rendement	7
1.3.2	Le compromis risque-rendement	8
1.3.3	Rendement Espéré	8
1.3.4	La Volatilité	9
1.3.5	Covariance	10
1.3.6	Coefficient de Corrélation	10
1.3.7	Diversification	11
1.3.8	Frontière Efficente	11
1.4	Illustration de la théorie moderne du portefeuille : diversification et réduction de la volatilité	14
1.5	Limites de la MPT	14
1.6	Les approches modernes : réponses aux limites de la MPT	15
1.7	Portefeuilles à risques multiples : de la finance à la gestion stratégique des actifs réels	16
1.8	Conclusion	18
2	Gestion de portefeuille dans le secteur de l'assurance	19
2.1	Le Portefeuille d'Assurance à Double Risque	19
2.2	De la théorie de Markowitz à l'ALM (Asset Liability Management)	19
2.3	Dynamique Actif-Passif (DAF / ALM)	20
2.3.1	Métriques	20
2.4	Cadre d'étude	21
2.5	Méthodologie : Estimation du Passif Agrégé L_t	21
2.5.1	Modélisation de la Fréquence (N_t)	21
2.5.2	Modélisation de la Sévérité (X_i)	22
2.5.3	Passif Agrégé Final	22
3	Présentation des données et Analyse descriptive	23
3.1	Présentation du jeu de données	23
3.2	Analyse descriptive	24
3.2.1	Variables quantitatives	24

3.2.2	Variables qualitatives	27
4	Résultats des modèles	29
4.1	Résultats et Sélection du Modèle de Fréquence	29
4.1.1	Évaluation du Modèle de Poisson	30
4.1.2	Comparaison Binomial Négatif versus Hurdle	30
4.1.3	Conclusion sur le Choix du Modèle	30
4.2	Interprétation du Modèle de Fréquence	31
4.2.1	Conclusion Définitive sur le Choix du Modèle de Fréquence	31
4.2.2	Synthèse des Coefficients du Modèle de Fréquence Retenu	32
4.3	Modélisation de la Sévérité des Sinistres (X_i)	33
4.3.1	Interprétation des Coefficients du Modèle de Sévérité	34
4.3.2	Limitation du Modèle : Le \mathbf{R}^2 faible	34
	Conclusion	35
	Bibliographie	36

Table des figures

1.1	Risque et rendement des actifs, portefeuilles simulés et frontière efficiente de Markowitz	12
1.2	Risque et rendement des actifs et portefeuilles simulés, avec la frontière efficiente et la CML.	13

Chapitre 1

Fondements de la gestion de portefeuille

1.1 Introduction

La crise économique financière de 1929 a profondément changé la vision traditionnelle des marchés et de l'investissement. Avant cet événement, la Bourse était surtout perçue comme un terrain de spéculation, guidée par l'intuition, la confiance et parfois l'euphorie collective, plutôt que comme un domaine soumis à une analyse rigoureuse. Les décisions d'investissement se fondaient principalement sur l'espoir de gains futurs, sans vraiment prendre en compte le risque ni la corrélation entre les actifs. L'ampleur de cette crise a provoqué une propagation à l'économie réelle, déclenchant un effet domino qui a conduit à un ralentissement économique puis à une récession mondiale. Une des causes majeures a été la mauvaise estimation des risques liés aux titres adossés aux crédits hypothécaires américains, notamment par les agences de notation, ainsi que leur multiplication excessive. Deux décennies plus tard, les modèles financiers comme celui de Black et Scholes, utilisés pour le trading d'options, ont été critiqués pour ne pas prendre en compte certains risques résiduels, ce qui a contribué au krach boursier de 1987. Cela souligne l'importance cruciale d'une estimation précise du risque en finance et les conséquences dramatiques d'une mauvaise gestion et d'une diversification insuffisante. De nombreuses entreprises et investisseurs trop concentrés sur certains secteurs ont alors subi des pertes importantes. Ce tournant a conduit à une remise en question profonde des pratiques financières, ouvrant la voie à des méthodes plus rigoureuses pour évaluer et gérer les investissements. Dans ce contexte économique incertain, marqué par la volatilité des marchés et la complexité croissante des instruments financiers, la gestion de portefeuille devient essentielle dans la prise de décision. Les investisseurs, qu'ils soient particuliers, institutionnels ou compagnies d'assurance, cherchent à optimiser le couple rendement-risque pour allouer efficacement leurs ressources. C'est dans cette optique que la théorie moderne du portefeuille, développée par Harry Markowitz dans les années 1950, a marqué un tournant majeur dans la finance contemporaine.

1.2 Présentation de la théorie de Markowitz

1.2.1 Revue de Littérature

La théorie moderne du portefeuille (Modern Portfolio Theory, MPT) a été formalisée en 1952 par Harry Markowitz dans le cadre de sa thèse de doctorat en statistique. L'apport majeur de son modèle réside dans la modélisation de l'effet de la diversification, selon le nombre d'actifs détenus et leurs covariances respectives (Megginson, 1996, p. 325). Ses travaux, publiés sous le titre « Portfolio Selection » dans *The Journal of Finance*, ont été approfondis ultérieurement dans son ouvrage « Portfolio Selection : Efficient Diversification » (1959). Une trentaine d'années plus tard, Markowitz s'est vu décerner le prix Nobel pour ses contributions fondatrices à la théorie financière. En 1958, l'économiste James Tobin a enrichi cette théorie dans son article « Liquidity Preference as Behavior Toward Risk », publié dans *Review of Economic Studies*, en introduisant les notions de frontière efficiente et de ligne du marché des capitaux. S'appuyant sur les travaux de Markowitz, Tobin a montré que, indépendamment de leur aversion au risque, les investisseurs rationnels détiennent les actifs risqués dans les mêmes proportions, à condition d'avoir des anticipations homogènes (Megginson, 1996, citant Tobin, 1958). Leurs portefeuilles ne diffèrent donc que par la part allouée aux actifs sans risque. Parallèlement, William Sharpe, John Lintner et Jan Mossin ont développé, de manière indépendante, le Modèle d'évaluation des actifs financiers (Capital Asset Pricing Model, CAPM), directement issu des fondements posés par Markowitz et Tobin (Megginson, 1996, p. 325). Ce modèle a constitué une avancée clé dans la compréhension de l'équilibre des marchés, en permettant d'évaluer un actif en fonction de son risque systématique. En 1964, Sharpe a formalisé le CAPM en s'appuyant sur la frontière efficiente et la ligne du marché des capitaux, une contribution qui lui a valu le prix Nobel d'économie. L'année suivante, Lintner (1965) a reformulé le modèle du point de vue d'une entreprise émettrice d'actions, tandis que Mossin (1966) l'a redérivé indépendamment en introduisant des fonctions d'utilité quadratiques (Megginson, 1996, p. 327). Depuis ces contributions initiales, la MPT a fait l'objet de nombreuses extensions. Ce travail s'intéresse à une lacune persistante dans cette littérature : un certain écart par rapport à la simplicité, ainsi qu'une sous-utilisation des progrès réalisés en finance et en technologie. Il s'inscrit également dans la lignée des travaux du professeur Simon Benninga (Wharton School), auteur de *Principles of Finance with Excel* (2006), qui milite pour une approche pédagogique et accessible dans l'enseignement des concepts mathématiques de la MPT.

1.2.2 Origine et fondements de la théorie moderne du portefeuille

Dans la librairie de la Graduate Business School de l'Université de Chicago, un jeune homme était absorbé par la lecture d'un ouvrage de référence de son époque : *The Theory of Investment Value* de John Burr Williams (1938). À la fin de sa lecture, il se posa cette question : « Pourquoi L'auteur traite les flux futurs d'une entreprise comme s'ils étaient certains, alors qu'en réalité ils sont incertains. Cette incertitude est précisément ce qui rend les actions risquées. » Ce jeune homme, Harry Markowitz, venait ainsi de cerner un des problèmes fondamentaux de la finance : le risque. Il décida alors de consacrer sa thèse à développer une méthode rigoureuse et mesurable pour l'analyser. En 1952, il bouleversa la manière de prendre des décisions d'investissement en introduisant une approche scientifique basée sur la diversification optimale. Avant lui, l'essentiel des analyses portaient uniquement sur la rentabilité individuelle des actifs, sans tenir compte des relations entre eux. Markowitz démontra que le risque d'un portefeuille dépend autant du risque spécifique à chaque actif que des corrélations entre leurs rendements. Cette idée novatrice donna naissance à la Théorie Moderne du Portefeuille (Modern Portfolio Theory – MPT), qui constitue désormais la pierre angulaire de la finance moderne et des méthodes quantitatives d'optimisation des actifs. Son modèle mathématique permet d'identifier les portefeuilles dits efficaces, qui offrent le meilleur compromis entre risque et rendement. Les travaux de Markowitz ont ouvert la voie aux modèles financiers majeurs tels que le Capital Asset Pricing Model (CAPM) et l'Arbitrage Pricing Theory (APT). Plus de soixante ans après leur publication, ses concepts continuent d'avoir une influence déterminante sur la recherche et la pratique en gestion d'actifs.

1.2.3 La Question Révolutionnaire de Markowitz

Avant l'avènement des travaux de Markowitz, la préoccupation principale des investisseurs consistait à identifier la meilleure action à acquérir, en se focalisant sur la valeur intrinsèque de chaque entreprise, ses perspectives bénéficiaires et la qualité de sa gestion. Markowitz a radicalement transformé cette approche en formulant une question innovante et plus pertinente : quelle combinaison optimale d'actifs détenir dans un portefeuille ? Il fut pionnier dans la conceptualisation du risque boursier comme une variable mesurable et quantifiable, pouvant être traitée avec rigueur mathématique. En assimilant chaque actif financier à une variable aléatoire, il permit de caractériser celui-ci à travers deux paramètres essentiels : son rendement espéré, représentant la récompense potentielle, et sa variance ou écart-type, traduisant son niveau de risque. L'innovation majeure de Markowitz réside dans le fait de ne plus considérer les actifs isolément, mais d'analyser les corrélations entre les actifs composant le portefeuille. Contrairement à la vision traditionnelle qui considérait un portefeuille comme la simple somme de ses actifs, il démontra que ce dernier doit être envisagé comme un système interconnecté. Cette interconnexion implique que les corrélations entre actifs jouent un rôle prépondérant dans l'évaluation du risque global du portefeuille. De manière contre-intuitive, un actif perçu comme très risqué individuellement peut contribuer à réduire le risque total du portefeuille s'il évolue en opposition avec d'autres actifs. Dès lors, la gestion patrimoniale efficace repose davantage sur les relations entre actifs que sur les caractéristiques individuelles de chaque composant. Cette avancée conceptuelle a profondément influencé les pratiques et théories modernes de gestion de portefeuille, marquant un tournant dans l'optimisation financière.

1.3 Concepts Fondamentaux

La Théorie Moderne du Portefeuille (MPT) comprend de nombreuses hypothèses concernant les marchés et les investisseurs. Certaines de ces hypothèses sont explicites, tandis que d'autres sont implicites. Pour développer la MPT, Markowitz s'est appuyé sur les hypothèses clés suivantes :

- Les investisseurs sont rationnels (ils cherchent à maximiser les rendements tout en minimisant le risque),
- Les investisseurs ne sont disposés à accepter un risque plus élevé que si celui-ci est compensé par des rendements attendus plus élevés,
- Les investisseurs reçoivent en temps voulu toutes les informations pertinentes liées à leur décision d'investissement,
- Les investisseurs peuvent emprunter ou prêter un montant illimité de capital à un taux d'intérêt sans risque,
- Les marchés sont parfaitement efficaces,
- Les marchés n'incluent pas de frais de transaction ou de taxes,
- Il est possible de sélectionner des titres dont la performance individuelle est indépendante des autres investissements du portefeuille.

Ces hypothèses fondamentales ont été largement contestées. De nombreuses critiques de la théorie seront présentées à la fin du Chapitre 1. Mais avant d'aborder ces critiques, il est essentiel de bien comprendre les trois concepts fondamentaux qui structurent la théorie de Markowitz : le rendement espéré, le risque (mesuré par la volatilité) et la corrélation/diversification.

1.3.1 Risque et Rendement

Le risque financier se définit comme l'écart observé entre les rendements effectivement obtenus et les rendements historiquement attendus sur une période donnée (Bofah, s.d. ; McClure, 2010). Toutefois, selon la théorie de la sélection de portefeuille développée par Harry Markowitz, « l'aspect essentiel du risque d'un actif ne réside pas dans le risque propre à chaque actif pris isolément, mais dans la contribution de chaque actif au risque global du portefeuille ». Ainsi, le risque associé à un titre peut être appréhendé sous deux angles distincts :

- Individuel, lorsque l'actif est considéré isolément ;
- Au niveau du portefeuille, lorsque l'actif fait partie d'un ensemble d'actifs.

Dans le cadre d'un portefeuille, le risque total d'un titre se décompose en deux composantes principales :

- Le risque systématique, également appelé risque de marché ou risque commun ;
- Le risque non systématique, aussi désigné comme risque spécifique ou risque diversifiable (Lowering Portfolio Risk, 2011).

La théorie moderne du portefeuille (MPT) affirme que ces deux formes de risque sont essentielles à tout portefeuille d'investissement. Le risque systématique correspond à un

risque de nature macroéconomique, affectant simultanément un grand nombre d'actifs, bien que dans des proportions variables. Les principaux déterminants de ce risque incluent les conditions économiques générales telles que l'inflation, les taux d'intérêt, le chômage, les variations des taux de change ou encore le niveau du produit national brut. Ces facteurs exercent une influence sur la quasi-totalité des titres du marché, ce qui rend le risque systématique impossible à éliminer par la diversification. À l'opposé, le risque non systématique est de nature microéconomique. Il résulte de facteurs propres à une entreprise ou à un secteur d'activité particulier et demeure indépendant des fluctuations globales du marché. À titre d'illustration, une modification impopulaire de la politique tarifaire de Netflix a conduit à une baisse de la satisfaction des abonnés, à une diminution des revenus et à une chute du cours de son action. Néanmoins, cet événement n'a eu qu'un impact marginal sur les grands indices boursiers (Dow Jones, S&P 500) et sur le secteur des médias et du divertissement. Il a tout de même profité à Blockbuster Video, qui a bénéficié de la perte de parts de marché de Netflix. D'autres exemples de risque non systématique incluent la dégradation de la note de crédit d'une entreprise, la publication d'informations négatives dans les médias, ou encore le déclenchement d'une grève affectant une société donnée (Helela, s.d.). Bien que ce type de risque puisse être atténué par une diversification adéquate du portefeuille (McClure, 2010), il ne peut être entièrement supprimé. En effet, les rendements des actifs financiers présentent une corrélation partielle, si bien que le risque non systématique subsiste, même dans les portefeuilles les plus diversifiés.

$$\sigma_i^2 = \beta_i^2 \sigma_m^2 + \sigma_{\epsilon_i}^2 \quad (1.1)$$

$\beta_i = \frac{\text{Cov}(R_i, R_m)}{\sigma_m^2}$ sensibilité au marché, σ_m^2 variance du marché, et $\sigma_{\epsilon_i}^2$ variance spécifique.

1.3.2 Le compromis risque-rendement

Le principe du compromis risque-rendement, central dans la théorie de Markowitz, établit qu'un niveau de risque plus élevé doit être compensé par une espérance de rendement supérieure. Le risque, défini comme l'écart possible entre rendement attendu et réalisé, est mesuré par la volatilité (écart type). Les investisseurs exigent une prime de risque pour accepter une incertitude accrue, mais seule la part non diversifiable du risque est rémunérée. Ainsi, la prise de risque ne garantit pas un rendement plus élevé ; elle en augmente seulement la probabilité potentielle. Dans ce cadre, la volatilité sert de mesure principale du risque à travers la variance, la covariance et la corrélation entre actifs.

1.3.3 Rendement Espéré

Le rendement espéré correspond à la moyenne pondérée des rendements potentiels d'un actif, servant de base à l'élaboration de portefeuilles selon le modèle de Markowitz. Bien que souvent dérivé des performances historiques, son estimation reste limitée par l'incertitude liée à la période d'observation. Seules des données historiques étendues et consistantes offrent un indicateur crédible pour anticiper le rendement futur.

Le rendement espéré d'un actif i est donné par :

$$E[R_i] = \sum_{k=1}^n p_k R_{i,k} \quad (1.2)$$

où $R_{i,k}$ est le rendement possible dans le scénario k et p_k sa probabilité.

Pour un portefeuille de N actifs avec des poids w_i :

$$E[R_p] = \sum_{i=1}^N w_i E[R_i] \quad (1.3)$$

1.3.4 La Volatilité

La volatilité est une mesure fondamentale du risque financier, reflétant l'ampleur des fluctuations des rendements d'un actif par rapport à sa moyenne. Elle constitue un indicateur clé pour évaluer l'incertitude associée à un investissement. La volatilité peut être quantifiée principalement par la variance et l'écart type. La variance correspond à la moyenne des écarts au carré entre les rendements observés et le rendement attendu, offrant ainsi une mesure statistique de la dispersion des rendements. Dans le cadre d'un portefeuille, la variance prend en compte non seulement les fluctuations individuelles des actifs, mais également leurs interactions, mesurées par les covariances et les corrélations. La diversification permet ainsi de réduire le risque global, car les pertes potentielles d'un actif peuvent être partiellement compensées par les gains d'autres actifs. L'écart type, défini comme la racine carrée de la variance, constitue une mesure directement interprétable de la volatilité. Dans le modèle de sélection de portefeuille de Markowitz, il est supposé que les investisseurs prennent leurs décisions en fonction du rendement attendu et du niveau de risque associé. Un écart type plus élevé indique un risque plus important, ce qui implique qu'un rendement potentiel supérieur est requis pour compenser ce risque. Ainsi, la variance et l'écart type offrent des outils statistiques essentiels pour quantifier la volatilité des actifs financiers et des portefeuilles, permettant aux investisseurs d'évaluer de manière rigoureuse le risque associé à leurs choix d'investissement.

La variance mesure la dispersion des rendements autour de la moyenne :

$$\sigma_i = \sqrt{\sigma_i^2} \quad (1.4)$$

Pour un portefeuille :

$$\sigma_p^2 = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \text{Cov}(R_i, R_j) \quad (1.5)$$

1.3.5 Covariance

La variance et l'écart type permettent de mesurer la variabilité individuelle des rendements d'un titre. Toutefois, pour évaluer la relation entre les rendements de deux titres distincts, il est nécessaire de recourir à la covariance ou à la corrélation. Ces deux mesures statistiques permettent d'apprécier la manière dont deux variables aléatoires sont liées. La covariance quantifie spécifiquement l'interdépendance des rendements de deux actifs financiers. Une covariance positive indique que les rendements tendent à évoluer dans le même sens, tandis qu'une covariance négative traduit un mouvement inverse. Si les rendements ne présentent aucune relation, la covariance est nulle. Dans le cadre de la théorie moderne du portefeuille, Markowitz souligne l'importance de la covariance pour la diversification : « Il est nécessaire d'éviter d'investir dans des titres présentant des covariances élevées entre eux » (Markowitz, 1952, p. 89). La prise en compte des covariances permet ainsi de construire des portefeuilles optimisés, minimisant le risque global tout en conservant un rendement attendu satisfaisant.

La covariance entre deux actifs i et j :

$$\text{Cov}(R_i, R_j) = \sum_{k=1}^n p_k (R_{i,k} - E[R_i])(R_{j,k} - E[R_j]) \quad (1.6)$$

1.3.6 Coefficient de Corrélation

Le coefficient de corrélation constitue une mesure essentielle pour évaluer la relation entre les rendements de deux actifs financiers. Il permet de déterminer dans quelle mesure les variations d'un actif sont liées à celles d'un autre, tout en standardisant les valeurs issues de la covariance. Dans la théorie moderne du portefeuille (MPT), le coefficient de corrélation est utilisé pour quantifier l'effet de diversification sur le risque global d'un portefeuille. Il se calcule en divisant la covariance des rendements de deux titres par le produit de leurs écarts types. Une corrélation positive indique que les rendements évoluent dans le même sens, une corrélation négative qu'ils évoluent de manière opposée, et une corrélation nulle qu'ils sont indépendants. L'efficacité de la diversification dépend du niveau de corrélation entre les actifs et de leur pondération dans le portefeuille (Wecker, n.d.). La présence d'actifs faiblement corrélés ou non corrélés contribue à réduire significativement le risque global. Les corrélations « imparfaites », comprises entre -1 et +1, favorisent généralement une diminution du risque du portefeuille, tandis que des corrélations élevées limitent l'effet de diversification (Gibson, 2004 ; Hight, 2010). Toutefois, la corrélation peut exister même entre des actifs sans lien sectoriel ou économique apparent, ce qui nécessite une analyse prudente lors de la sélection des investissements. Ainsi, le coefficient de corrélation constitue un indicateur central pour la gestion du risque et l'optimisation des portefeuilles, en permettant de quantifier et d'exploiter l'effet de diversification.

$$\rho_{i,j} = \frac{\text{Cov}(R_i, R_j)}{\sigma_i \sigma_j}, \quad -1 \leq \rho_{i,j} \leq 1 \quad (1.7)$$

1.3.7 Diversification

La diversification constitue un principe central de la théorie moderne du portefeuille et de la théorie de Markowitz, visant à répartir les investissements entre différents titres, secteurs et classes d'actifs afin de réduire le risque global. Elle repose sur le principe de ne pas « mettre tous ses œufs dans le même panier », permettant de limiter les pertes potentielles tout en optimisant le rendement. L'effet de diversification découle de la corrélation imparfaite entre les actifs : plus les actifs sont faiblement corrélés, plus le risque global du portefeuille est réduit sans compromettre le rendement. Toutefois, la diversification ne peut éliminer totalement le risque. Le risque non systématique, spécifique à certaines entreprises ou secteurs, peut être réduit par la diversification, tandis que le risque systématique, lié à des facteurs macroéconomiques ou géopolitiques, demeure inévitable et affecte l'ensemble du marché. Ainsi, la diversification permet d'améliorer l'équilibre rendement/risque, mais ne supprime jamais entièrement l'exposition au risque.

La variance d'un portefeuille montre l'effet de diversification :

$$\sigma_p^2 = \sum_i w_i^2 \sigma_i^2 + \sum_i \sum_{j \neq i} w_i w_j \sigma_i \sigma_j \rho_{i,j} \quad (1.8)$$

1.3.8 Frontière Efficiente

La frontière efficiente de Markowitz, élément clé de la MPT (Efficient Frontier/Money Terms, n.d.), représente l'ensemble des portefeuilles offrant le rendement attendu maximal pour un niveau de risque donné. Cette relation entre le rendement attendu d'un portefeuille et sa volatilité est généralement représentée graphiquement sous la forme d'une courbe comparant risque et rendement. Les portefeuilles situés sur cette frontière optimisent le compromis rendement/risque, offrant la meilleure combinaison possible entre rendement attendu et risque pour un niveau de risque déterminé. La relation entre les actifs au sein d'un portefeuille joue un rôle fondamental dans la construction de la frontière efficiente. Certains titres peuvent évoluer dans le même sens, tandis que d'autres se déplacent en sens opposé. Plus les actifs sont faiblement ou négativement corrélés, plus la covariance réduit la volatilité globale du portefeuille, diminuant ainsi le risque. L'un des principaux enseignements de la théorie de Markowitz réside dans les avantages de la diversification. En combinant des actifs dont les mouvements ne sont pas parfaitement corrélés, il est possible d'augmenter le rendement attendu d'un portefeuille sans accroître le risque. La théorie suggère que des investisseurs rationnels chercheront à constituer des portefeuilles situés sur la frontière efficiente afin de maximiser leurs rendements pour un risque minimal.

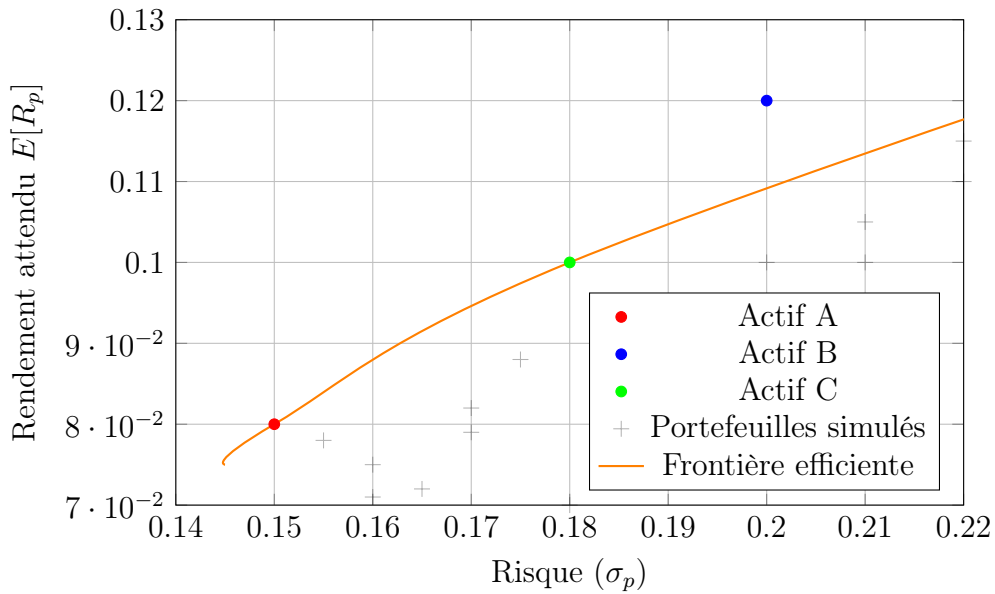


FIGURE 1.1 – Risque et rendement des actifs, portefeuilles simulés et frontière efficiente de Markowitz

Explications :

Le graphique illustre le risque et le rendement des actifs individuels, des portefeuilles simulés ainsi que la frontière efficiente de Markowitz. L'axe horizontal représente le risque du portefeuille, mesuré par l'écart-type (σ_p), tandis que l'axe vertical indique le rendement attendu ($E[R_p]$).

Les points rouges, verts et bleus correspondent aux actifs A, C et B, respectivement, tandis que les petites croix noires représentent différents portefeuilles simulés, c'est-à-dire des combinaisons possibles de ces actifs. La courbe orange lisse trace la frontière efficiente, qui regroupe l'ensemble des portefeuilles optimaux pour chaque niveau de risque, elle identifie le rendement attendu maximal.

Les portefeuilles situés en dessous de cette courbe sont considérés comme inefficaces, dans la mesure où ils présentent soit un risque plus élevé pour un rendement donné, soit un rendement inférieur pour un niveau de risque équivalent. Ce graphique met ainsi en évidence l'intérêt de la diversification en combinant les actifs de manière optimale, il est possible de constituer des portefeuilles offrant le meilleur compromis entre risque et rendement, conformément aux principes de la théorie moderne du portefeuille.

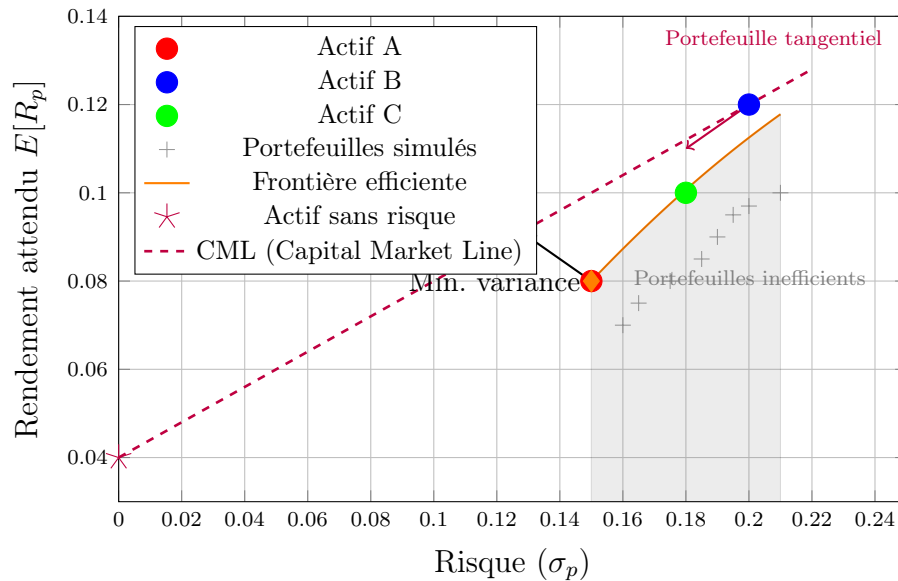


FIGURE 1.2 – Risque et rendement des actifs et portefeuilles simulés, avec la frontière efficiente et la CML.

Explication :

- **Actifs individuels (A, B, C) :** Chaque point coloré représente un actif pris isolément, avec son rendement attendu $E[R]$ et son risque σ . Chaque actif a un compromis rendement/risque propre.
- **Portefeuilles simulés :** Les points gris illustrent des portefeuilles composés de combinaisons des actifs. La diversification permet de réduire le risque pour un même rendement ou d'augmenter le rendement pour un même niveau de risque.
- **Frontière efficiente :** La courbe orange montre les portefeuilles offrant le meilleur rendement pour chaque niveau de risque. Les portefeuilles situés sous la frontière (zone grisée) sont inefficients.
- **Portefeuille de variance minimale :** Point orange en losange à l'extrémité gauche de la frontière. Il représente le portefeuille le moins risqué tout en restant efficient.
- **Actif sans risque :** Étoile violette, par exemple un placement sûr. Il sert de base pour la CML.
- **Capital Market Line (CML) :** Ligne violette en pointillé reliant l'actif sans risque au portefeuille tangentiel sur la frontière efficiente. Elle représente le meilleur compromis rendement/risque pour les investisseurs pouvant emprunter ou prêter au taux sans risque.
- **Portefeuille tangentiel :** Point sur la frontière où la CML est tangente, offrant le meilleur ratio rendement/risque parmi tous les portefeuilles combinant actifs risqués et sans risque.
- **Zone grisée :** Les portefeuilles situés sous la frontière sont inefficients et doivent être évités.

1.4 Illustration de la théorie moderne du portefeuille : diversification et réduction de la volatilité

Scénario Météo	Glacier	Parapluies	Portefeuille 50/50
Soleil intense	+30%	-20%	+5%
Pluie diluvienne	-20%	+30%	+5%
Moyenne	+5%	+5%	+5%
Volatilité (Écart-type)	Très élevée	Très élevée	Nulle

Ce tableau illustre de manière simple le principe fondamental de la MPT. Il compare les rendements de deux activités très différentes dont un glacier et un vendeur de parapluies selon divers scénarios météorologiques. Pris séparément, le glacier et le vendeur de parapluies connaissent des rendements très volatils : chacun dépend fortement des conditions météorologiques. En revanche, lorsque l'on combine les deux activités dans un portefeuille équilibré (50/50), les variations extrêmes s'annulent. Le rendement moyen reste positif (+5 %) et la volatilité devient nulle. Cette démonstration met en évidence que ce n'est pas seulement la performance individuelle d'un actif qui compte, mais la manière dont il interagit avec les autres. La diversification permet de réduire le risque global sans sacrifier le rendement attendu. C'est précisément ce mécanisme qui fonde la frontière efficiente et qui constitue la rupture introduite par Markowitz dans les années 1950.

1.5 Limites de la MPT

Bien que la théorie moderne du portefeuille (MPT) représente une avancée majeure dans la gestion financière, elle présente plusieurs limites en raison de l'écart entre ses hypothèses et la réalité des marchés. Premièrement, la MPT repose sur l'idée que les investisseurs sont rationnels et cherchent à maximiser le rendement tout en minimisant le risque. Or, les comportements observés montrent souvent l'inverse : les investisseurs peuvent suivre la foule, privilégier des secteurs « à la mode » et provoquer des cycles de boom et de krach spéculatif. Deuxièmement, l'hypothèse selon laquelle un risque plus élevé doit être compensé par un rendement supérieur est parfois contredite. Certains instruments considérés comme risqués peuvent être utilisés pour réduire le risque global, sans augmentation correspondante du rendement attendu, et les préférences individuelles des investisseurs peuvent primer sur la simple distribution des rendements. Troisièmement, la théorie suppose une information parfaite, accessible à tous en temps réel. En réalité, les marchés souffrent d'asymétrie d'information et d'avantages informationnels pour certains acteurs, ce qui peut conduire à des acquisitions en dessous de la valeur réelle des actifs. Quatrièmement, la MPT postule un accès illimité au capital à un taux sans risque, ce qui est irréaliste, car les capacités d'emprunt sont limitées et seuls certains acteurs, comme les gouvernements, peuvent bénéficier de taux quasi sans risque (Morien, n.d.). Cinquièmement, la théorie repose sur des marchés parfaitement efficaces. Or, les marchés réels sont affectés par des externalités, des asymétries d'information et des crises répétées,

démontrant que l'efficience parfaite est rarement atteinte. Enfin, la MPT ne prend pas en compte les coûts de transaction, la fiscalité, ni la corrélation dynamique des actifs, qui peut varier en période de stress, remettant en question l'indépendance supposée des investissements. D'autres critiques soulignent que la notion d'actif totalement sans risque n'existe pas et que la MPT ne considère pas les facteurs environnementaux, sociaux ou stratégiques dans l'allocation de portefeuille. En somme, malgré sa rigueur théorique et sa contribution à l'optimisation des portefeuilles, la MPT reste limitée dans sa capacité à représenter fidèlement les marchés financiers réels.

1.6 Les approches modernes : réponses aux limites de la MPT

Les limites identifiées dans la section précédente ont stimulé le développement d'approches alternatives qui étendent ou amendent le cadre markowitzien sans remettre en cause sa pertinence fondamentale :

La Post-Modern Portfolio Theory (PMPT) répond directement à la critique de la symétrie risque-rendement en se concentrant exclusivement sur le risque de baisse. Alors que Markowitz utilise la volatilité totale, supposant que les gains et les pertes sont également indésirables, la PMPT introduit des mesures unilatérales comme le downside risk, la semi-variance ou la Value at Risk, reconnaissant que les investisseurs ne craignent que les pertes, pas les gains inattendus. Cette approche est particulièrement pertinente pour les investisseurs averses au risque asymétrique.

Le modèle Black-Litterman répond au problème de la sensibilité aux inputs en combinant les équilibres de marché implicites avec les vues subjectives du gestionnaire. Cette méthode stabilise le processus d'optimisation et évite les allocations extrêmes générées par les estimations purement historiques, qui conduisent souvent à des portefeuilles irréalistes en pratique.

L'approche Risk Parity émerge comme une réponse aux échecs de la diversification traditionnelle lors des crises. En allouant le risque plutôt que le capital, elle offre une résilience quand les corrélations s'effondrent, comme observé en 2008, où les actifs supposés décorrélés ont chuté simultanément. Cette approche reconnaît que la diversification basée uniquement sur les montants investis ignore les contributions hétérogènes au risque total.

Ces approches, complétées par les enseignements de la finance comportementale sur les biais cognitifs et les techniques d'optimisation face à l'incertitude des paramètres, forment aujourd'hui un corpus étendu qui enrichit le cadre markowitzien sans le rendre obsolète.

1.7 Portefeuilles à risques multiples : de la finance à la gestion stratégique des actifs réels

La théorie moderne du portefeuille, initialement développée pour l'allocation d'actifs financiers cotés, repose sur l'optimisation du compromis entre espérance de rendement et risque (mesuré par la variance), grâce à la diversification c'est-à-dire à la sélection d'actifs présentant des rendements faiblement corrélés. Si cette approche a longtemps été cantonnée aux marchés financiers, une littérature croissante (Dixit & Pindyck, 1994 ; Teece, 2000 ; Trigeorgis, 1996) souligne sa pertinence dans des contextes où les « actifs » ne sont pas monétaires ou négociables, mais relèvent de l'opérationnel, du technologique ou du stratégique.

La distinction essentielle réside non dans la nature des outils mathématiques mobilisés, mais dans la définition opérationnelle des variables clés. Dans un portefeuille financier :

- le rendement est un flux monétaire observable (dividende, plus-value),
- le risque est la volatilité statistique de ce flux,
- les actifs sont des titres échangeables sur des marchés liquides.

À l'inverse, dans un portefeuille non financier, qu'il s'agisse de contrats d'assurance, de concessions minières ou de brevets :

- le « rendement » désigne une valeur économique, technique ou stratégique prospectivement attendue (ex. : production nette, chiffre d'affaires, couverture d'un passif),
- le « risque » renvoie à la variabilité potentielle de ce résultat, induite par des aléas exogènes (climatiques, réglementaires, technologiques) ou endogènes (comportementaux, opérationnels),
- la covariance entre actifs traduit le degré de dépendance conjointe à des chocs structurels communs (ex. : une sécheresse affectant plusieurs cultures, une réforme légale impactant un domaine technologique entier).

Cette transposition conceptuelle suppose que les décideurs puissent estimer, même imparfaitement, l'espérance, la dispersion et les interdépendances des résultats associés à chaque composante du portefeuille ce qui implique souvent le recours à des modèles probabilistes, des expertises sectorielles ou des simulations stochastiques. Elle ne vise pas à reproduire mécaniquement la frontière efficiente de Markowitz, mais à structurer la réflexion stratégique autour de trois principes fondamentaux :

1. expliciter les sources multiples de risque,
2. évaluer leur degré d'interdépendance,
3. allouer les ressources de manière à réduire la vulnérabilité systémique du portefeuille global.

Les exemples qui suivent illustrent cette démarche dans sept contextes représentatifs, où la logique de diversification sert à stabiliser les résultats économiques ou opérationnels face à l'incertitude, bien au-delà des frontières de la finance traditionnelle.

Portefeuille d'assurance automobile

Le portefeuille se compose de contrats caractérisés par deux sources de risque principales : (i) la fréquence des sinistres (aléa de survenance), et (ii) la sévérité (aléa de coût). La rentabilité technique dépend de l'espérance et de la variance de la charge sinistres totale. Une diversification par segments de risque (profil d'assuré, zone géographique, usage du véhicule) permet de réduire la covariance des charges entre sous-portefeuilles, conformément au principe de mutualisation (Borch, 1962). La tarification différenciée (ex-ante) et la réassurance (ex-post) constituent des mécanismes complémentaires d'optimisation du couple espérance-variance.

Portefeuille de concessions minières

Chaque concession est un actif réel à rendement incertain, soumis à des risques séquentiels : exploration (succès/fiasco), évaluation des réserves (tonnage, teneur), exploitation (coûts, réglementation), et marché (prix des métaux). L'allocation inter-concessions vise à minimiser la variance du cash-flow global en combinant des projets à différents stades de maturité et localisés dans des bassins géologiques ou juridiques décorrélés (Trigeorgis, 1996). Les options réelles (abandon, expansion) renforcent la flexibilité stratégique, élargissant la frontière efficiente au-delà de la MPT statique.

Portefeuille de brevets

Un brevet constitue un actif immatériel dont la valeur dépend de la probabilité d'exploitation effective (risque technologique/commercial) et de la variabilité des flux futurs (licences, redevances). La diversification s'opère selon trois axes : (i) domaine technologique (réduction de la corrélation face aux ruptures), (ii) secteur d'application (atténuation du risque marché), et (iii) géographie (différenciation des régimes de protection). L'évaluation probabiliste (ex. : méthode risk-adjusted net present value) permet de hiérarchiser les actifs selon leur rapport espérance-risque (Reitzig, 2004).

Portefeuille éditorial

Les projets littéraires présentent une asymétrie forte : faible coût initial, mais distribution bimodale des ventes (« longue traîne » avec quelques succès exceptionnels). Le risque principal réside dans la difficulté à prédire la réception du public. La diversification par genre, notoriété de l'auteur, format (papier/numérique), ou segment de marché permet de stabiliser les revenus du catalogue et de lisser la volatilité du résultat opérationnel, conformément à la logique de portefeuille (Bhattacharya et al., 1998).

Portefeuille de cultures agricoles

Chaque culture est un actif biophysique dont le rendement économique est aléatoire, soumis à des risques climatiques, phytosanitaires et de marché. La covariance entre cultures dépend de leur sensibilité commune à des facteurs environnementaux (ex. : déficit hydrique). La diversification interculturelle (ex. : céréales et légumineuses) et l'adoption de pratiques agroécologiques (rotation, couverts) réduisent la variance du revenu global (Hazzell, 1979). L'optimisation sous contrainte (surface, rotation obligatoire, main-d'œuvre) relève de la programmation stochastique.

Portefeuille d’infrastructures énergétiques renouvelables

Les projets d’énergie renouvelable présentent des profils de risque–rendement différenciés selon la technologie (éolien, solaire, hydro), la localisation (ressource naturelle) et le stade de développement. La diversification géographique (décorrélation des régimes climatiques) et technologique (complémentarité saisonnière : éolien hivernal et solaire estival) permet de réduire la variabilité de la production et donc des revenus, en s’approchant d’un « portefeuille bas-carbone efficient » (Kumar et al., 2021).

Portefeuille d’actifs des caisses de retraite

Ici, la MPT est intégrée dans un cadre de gestion actif–passif (ALM). L’objectif n’est plus de minimiser la variance des rendements, mais le risque de déséquilibre actuariel (incapacité à couvrir les engagements futurs). La diversification vise alors à aligner la durée, l’indexation et la volatilité des actifs sur celles des passifs. Les actifs réels (infrastructures, immobilier) sont privilégiés pour leur faible corrélation avec les actifs traditionnels et leur capacité à générer des flux longs et indexés, réduisant ainsi le risque de taux et d’inflation (Sharpe, 2002). Les contraintes réglementaires (ex : Solvabilité II) imposent cependant des déviations contrôlées par rapport à la frontière efficiente théorique.

1.8 Conclusion

la théorie moderne du portefeuille a révolutionné la gestion financière. Elle donne aux investisseurs un cadre scientifique clair pour optimiser le couple rendement–risque et construire des portefeuilles plus robustes. Mais pour être pleinement efficace, elle doit être adaptée à la réalité des marchés, prendre en compte le comportement humain et intégrer la possibilité d’événements exceptionnels.

Chapitre 2

Gestion de portefeuille dans le secteur de l'assurance

2.1 Le Portefeuille d'Assurance à Double Risque

Nous avons pu voir dans les sections précédentes que la notion de portefeuille, établie par la Théorie Moderne de Markowitz, vise à optimiser le couple rendement-risque par la diversification. Cependant, nous avons souligné que son application s'étend au-delà de la finance pure pour englober les portefeuilles dits non financiers.

Cette section a pour objectif de détailler la problématique de la gestion du portefeuille d'assurance, qui nécessite une approche de **Gestion Actif-Passif (ALM)**. Contrairement à un portefeuille financier pur, l'engagement de l'assureur est sujet à un **double risque stochastique** :

1. Le risque de **Fréquence** (le nombre de sinistres, N_t) ;
2. Le risque de **Sévérité** (le montant des sinistres, X_i).

2.2 De la théorie de Markowitz à l'ALM (Asset Liability Management)

Le but ici n'est plus de considérer uniquement l'Actif, mais aussi le Passif. L'ALM est l'optimisation du portefeuille afin d'assurer la solvabilité d'une compagnie d'assurance sur le long terme.

Les premiers travaux ont été introduits par Redington en 1952 en posant l'équation de base de l'ALM :

$$S_t = A_t - L_t$$

- S_t : Le Surplus (ou Impasse) à l'instant t , représentant la différence entre l'Actif et le Passif.

- A_t : L'Actif. On peut y retrouver des actifs "sûrs" (obligations, titres d'État...), des actifs plus risqués (comme des actions) et, pour finir, les reçus (primes) des diverses polices d'assurance.
- L_t : Le Passif est ici simplifié aux diverses charges (coûts) que doit supporter une compagnie d'assurance (comme les prestations/indemnités à payer aux assurés).

Les premiers travaux de Redington avaient pour objectifs de maximiser le surplus terminal de la compagnie d'assurance, mais aussi de réduire la probabilité de ruine $P(S_t < 0)$.

2.3 Dynamique Actif-Passif (DAF / ALM)

Pour simuler ce genre de portefeuille, il est intéressant d'utiliser des méthodes de simulation stochastique, comme la simulation de Monte-Carlo. Le but est, après avoir calculé le coût moyen que rencontre une compagnie d'assurance (L_t) et ses revenus moyens de polices (A_t), de jouer sur une pondération w sur les actifs risqués et sûrs afin de déterminer quel portefeuille est le plus optimal pour éviter la ruine et assurer la plus grande solvabilité.

La simulation de Monte-Carlo va, en fonction de ces paramètres, générer un grand nombre de scénarios économiques (50 000 par exemple) et tirer des métriques pour évaluer la performance du portefeuille.

2.3.1 Métriques

Une fois la simulation effectuée, il va en ressortir plusieurs indicateurs clés pour l'évaluation de la performance du portefeuille et du risque de l'entreprise :

- **Le Surplus Terminal Moyen ($E[S_T]$)** : C'est la valeur espérée du surplus ($S_t = A_t - L_t$) à l'horizon T . Le gestionnaire cherche à le maximiser.
- **La Probabilité de Ruine ($P(S_T < 0)$)** : C'est la fréquence des scénarios simulés où le Passif excède l'Actif à l'horizon T . L'objectif est de minimiser cette probabilité.
- **La Valeur à Risque (VaR) du Surplus** : Elle mesure la perte maximale du surplus à un niveau de confiance donné (par exemple, 99% ou 99.5%). La VaR à 99% représente la perte du surplus qui n'est dépassée que dans 1% des scénarios les plus défavorables.
- **L'Excédent de Perte Attendu (Expected Shortfall - ES)** : Aussi appelé C-VaR, elle mesure la perte moyenne dans le pire $x\%$ des cas (au-delà de la VaR). C'est une métrique plus prudente et largement utilisée pour le calcul des exigences en capital réglementaire (comme Solvabilité II).
- **La Sensibilité aux Facteurs de Risque (Greeks/Duration)** : L'analyse de l'impact des variations des taux d'intérêt, des marchés actions ou de l'inflation sur le surplus S_t .

2.4 Cadre d'étude

Pour notre étude, nous nous focaliserons uniquement sur le Passif, étant donné la difficulté d'obtenir des données réelles d'assurance complètes. Nous avons néanmoins pu mettre la main sur une base modélisant bien les problématiques du Passif pour une compagnie d'assurance automobile.

Le but de cette analyse sera double :

1. Comprendre les facteurs jouant sur la **fréquence des sinistres** et sur leur **coût** (sévérité).
2. Calculer l'espérance du Passif ($E[L_t]$) pour obtenir une estimation du montant moyen total (les charges) auquel cette compagnie doit faire face, en fonction de son portefeuille d'assurés.

2.5 Méthodologie : Estimation du Passif Agrégé L_t

Le Passif agrégé (Perte Totale) L_t est la somme des coûts de tous les sinistres, décomposée en Fréquence (N_t) et Sévérité (X_i) :

$$L_t = \sum_{i=1}^{N_t} X_i$$

Notre approche repose sur l'utilisation des **Modèles Linéaires Généralisés (GLM)** pour modéliser séparément l'espérance de la Fréquence et de la Sévérité, en fonction des caractéristiques des assurés.

2.5.1 Modélisation de la Fréquence (N_t)

La fréquence est modélisée à l'aide de Modèles de Comptage. L'objectif est d'expliquer le nombre de sinistres ($N_{\text{sinistres}}$) en fonction des covariables (X_i).

1. **Lois de Comptage Mises en Concurrence** : Nous testons et comparons la *Loi de Poisson*, la *Loi Binomiale Négative* (pour capturer la surdispersion) et le *Modèle Hurdle* (pour gérer l'excès de zéros). Un seul modèle sera retenu après évaluation des critères d'ajustement.
2. **Le Modèle GLM** :

$$E[N_{\text{sinistres}}|\mathbf{X}] = \alpha + \log(\text{Durée_Couverture}) + \sum_{i=1}^k \beta_i \cdot X_i$$

Le terme $\log(\text{Durée_Couverture})$ sert d'offset pour normaliser la fréquence en fonction du temps d'exposition au risque. Les coefficients β_i permettent de quantifier l'impact de chaque facteur de risque X_i sur la fréquence.

2.5.2 Modélisation de la Sévérité (X_i)

La sévérité, X_i (coût unitaire d'un sinistre), est une variable continue positive et souvent asymétrique.

1. **Loi Log-normale** : Nous utilisons la loi Log-normale pour modéliser la distribution des coûts unitaires, car elle est adaptée aux données fortement asymétriques et garantit des coûts positifs.
2. **Le Modèle GLM** : Similairement à la fréquence, un GLM avec une fonction de lien logarithmique est utilisé pour modéliser l'espérance du coût en fonction des facteurs de risque :

$$\log(E[X_i|\mathbf{X}]) = \gamma_0 + \sum_{i=1}^k \gamma_i \cdot X_i$$

Les coefficients γ_i mesurent ici l'impact des facteurs de risque sur le coût moyen du sinistre.

2.5.3 Passif Agrégé Final

L'estimation de l'espérance du Passif ($E[L_t]$) est obtenue par la multiplication des espérances des deux modèles :

$$E[L_t] \approx E[N_t] \times E[X_i]$$

Chapitre 3

Présentation des données et Analyse descriptive

3.1 Présentation du jeu de données

Le jeu de données utilisé dans cette étude est `freMTPL2`, issu du package `CASdatasets` développé par Christophe Dutang et Arthur Charpentier. Il comprend 677 991 polices d'assurance automobile en Responsabilité Civile en France. Deux tables principales composent cette base :

- `freMTPL2freq`, qui rassemble les facteurs de risque et le nombre de sinistres déclarés pour chaque police ;
- `freMTPL2sev`, qui contient les informations relatives au coût des sinistres.

Afin d'obtenir une vision complète de chaque contrat, ces deux tables ont été jointes à l'aide de l'identifiant unique `IDpol`. Cette fusion permet de disposer, pour chaque observation, à la fois du nombre de sinistres et du montant associé, ce qui est indispensable pour la mise en œuvre d'une approche de tarification fréquence-sévérité. L'objectif est en effet d'expliquer, d'une part, le nombre d'événements survenus et, d'autre part, leur coût moyen.

La base finale comporte 12 variables, sans aucune valeur manquante, ce qui élimine la nécessité d'effectuer une imputation préalable. Parmi ces variables, certaines sont quantitatives comme le nombre et le montant des sinistres, la durée de couverture, la puissance du véhicule, l'âge du conducteur, l'âge du véhicule, le coefficient bonus-malus ou encore la densité de population. D'autres sont qualitatives, telles que la marque du véhicule, le type de carburant, la zone géographique et la région d'habitation.

Les deux variables centrales de notre analyse sont le nombre de sinistres et le montant des sinistres, qui constituent le cœur de la modélisation fréquence-sévérité.









Pour plus de clarté, le tableau ci-dessous présente une description synthétique des variables utilisées.

	Nom de la variable	Type	Description
1	Marque_Veh	factor	Catégorie de la marque automobile, divisée en modalités symboliques. (ex: B1, B2..)
2	Carburant	factor	Type de carburant du véhicule.
3	Zone	factor	Zone d'habitation du conducteur : urbain ou rural
4	Région	factor	Région de l'assuré en France.
5	Nb_Sinistres	integer	Nombre de sinistres survenus pendant la durée de couverture du contrat d'assurance.
6	Bonus_Malus	integer	Coefficient de réduction-majoration du conducteur : 50-100 = bonus ; >100 = malus (50 : bonus max ; 230 : malus max).
7	Age_Conducteur	integer	Âge du conducteur, en années.
8	Age_Veh	integer	Âge du véhicule, en années.
9	Puissance_Veh	integer	Puissance du véhicule, comprise entre 4 et 15 chevaux fiscaux.
10	Densité_Pop	integer	Densité de population de la ville de résidence de l'assuré, en nombre d'habitants par km².
11	Montant_Sinistre	numeric	Coût total des réclamations payées par l'assurance pour un assuré pendant sa période de couverture, en euros.
12	Durée_Couverture	numeric	Durée pendant laquelle l'assurance couvre l'assuré (exposition au risque), exprimée en années.

3.2 Analyse descriptive

3.2.1 Variables quantitatives

Statistiques descriptives des variables quantitatives

Noms	Densité	Moyenne	Ecart-type	Min	Max
Nb_Sinistres		0.06	0.28	0	16.00
Montant_Sinistre		88.16	5810.51	0	4075400.56
Durée_Couverture		0.53	0.36	0	2.01
Puissance_Veh		6.45	2.05	4	15.00
Age_Veh		7.04	5.66	0	100.00
Age_Conducteur		45.50	14.14	18	100.00
Bonus_Malus		59.78	15.67	50	230.00
Densité_Pop		1794.05	3960.83	1	27000.00

On remarque dans ce tableau que le nombre de sinistres par police apparaît très faible : la moyenne s'élève à 0,06, avec un maximum de 16 sinistres, révélant la présence de quelques valeurs extrêmes. L'écart-type, lui aussi très faible, égal à 0,28, confirme que la quasi-totalité des assurés ne déclarent aucun sinistre ou, au plus, un seul.

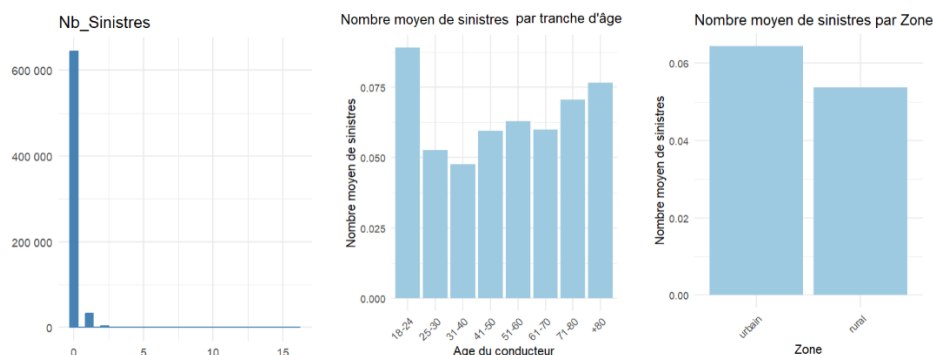
Le montant des sinistres est lui aussi très variable. Sa moyenne est de 88 euros, mais certains sinistres atteignent plus de 4 millions d'euros de dommages, illustrant une dis-

tribution fortement asymétrique, avec quelques valeurs très extrêmes qui font fortement augmenter la moyenne.

Le coefficient de bonus-malus affiche une moyenne d'environ 60, indiquant que la plupart des conducteurs disposent d'un niveau de tarification relativement favorable.

De manière générale, plusieurs variables se caractérisent par la présence de valeurs extrêmes susceptibles d'influencer la moyenne, tandis que certaines distributions se révèlent très concentrées autour de valeurs faibles, alors que d'autres présentent une dispersion beaucoup plus importante.

Répartition de la variable Nb_Sinistres

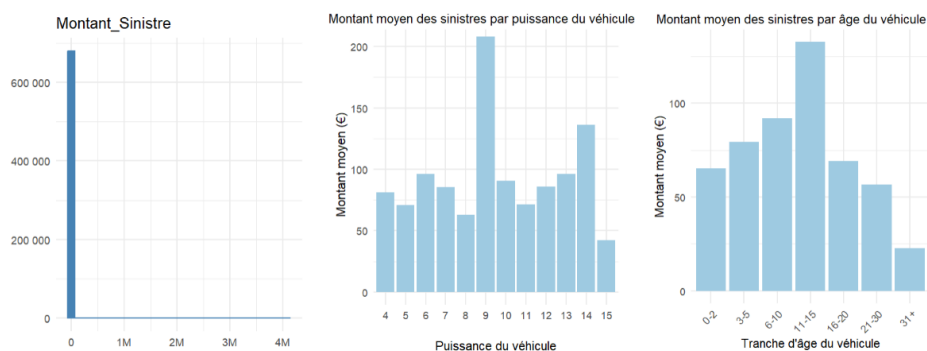


On observe très clairement, sur le premier graphique représentant la distribution simple du nombre de sinistres, une très forte concentration à 0, avec plus de 600 000 assurés n'ayant eu aucun sinistre, et une proportion très faible d'assurés ayant eu un sinistre.

Le deuxième graphique, qui montre le nombre moyen de sinistres en fonction de l'âge du conducteur, indique que les conducteurs ayant le plus d'accidents sont les jeunes de 18 à 24 ans, ainsi que les personnes âgées de plus de 80 ans, et on peut même élargir cette observation aux conducteurs de plus de 70 ans.

Le troisième graphique montre que les conducteurs habitant en zones urbaines ont en moyenne plus d'accidents que ceux vivant en zones rurales.

Répartition de la variable Montant_Sinistre

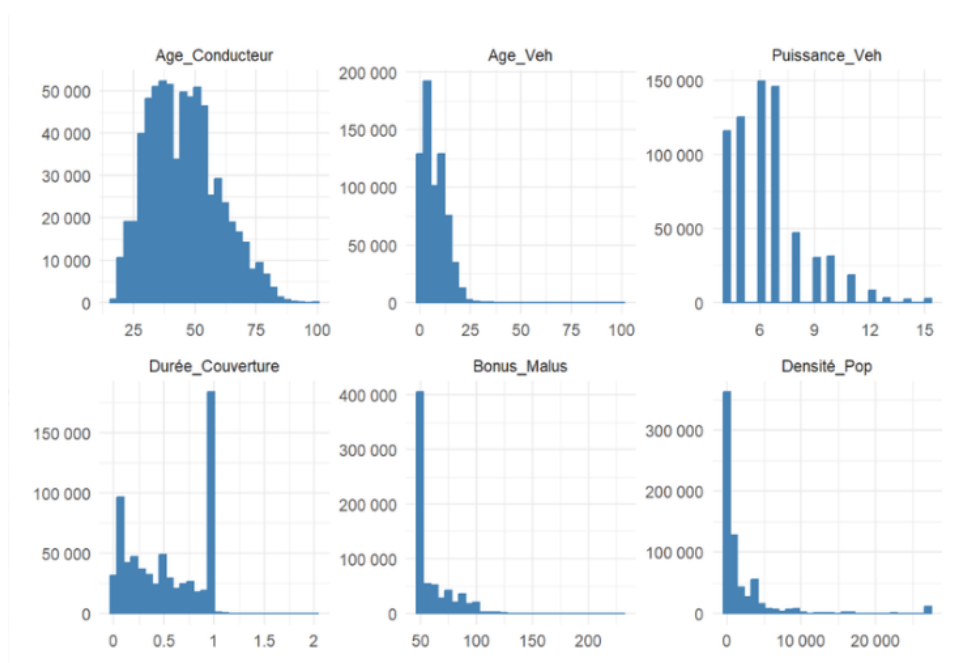


On remarque très clairement que la quasi-totalité des coûts des sinistres est égale à 0, avec toutefois quelques valeurs extrêmes pouvant atteindre jusqu'à 4 millions d'euros

pour certains sinistres. Ce premier graphique est cohérent avec le constat que nous avons fait précédemment concernant le faible nombre d'accidents survenus, puisque la majorité des polices déclarent aucun sinistre. Il est donc logique que la plupart des montants soient à 0.

Dans les deux autres graphiques, on remarque que les véhicules ayant une puissance de 9 chevaux fiscaux ainsi que ceux âgés de 11 à 15 ans entraînent en moyenne des coûts de dommages plus élevés.

Histogrammes des autres variables quantitatives



On constate que la majorité des assurés sont des adultes d'âge moyen, situés entre 30 et 55 ans. Les jeunes conducteurs ainsi que les conducteurs plus âgés sont beaucoup moins représentés.

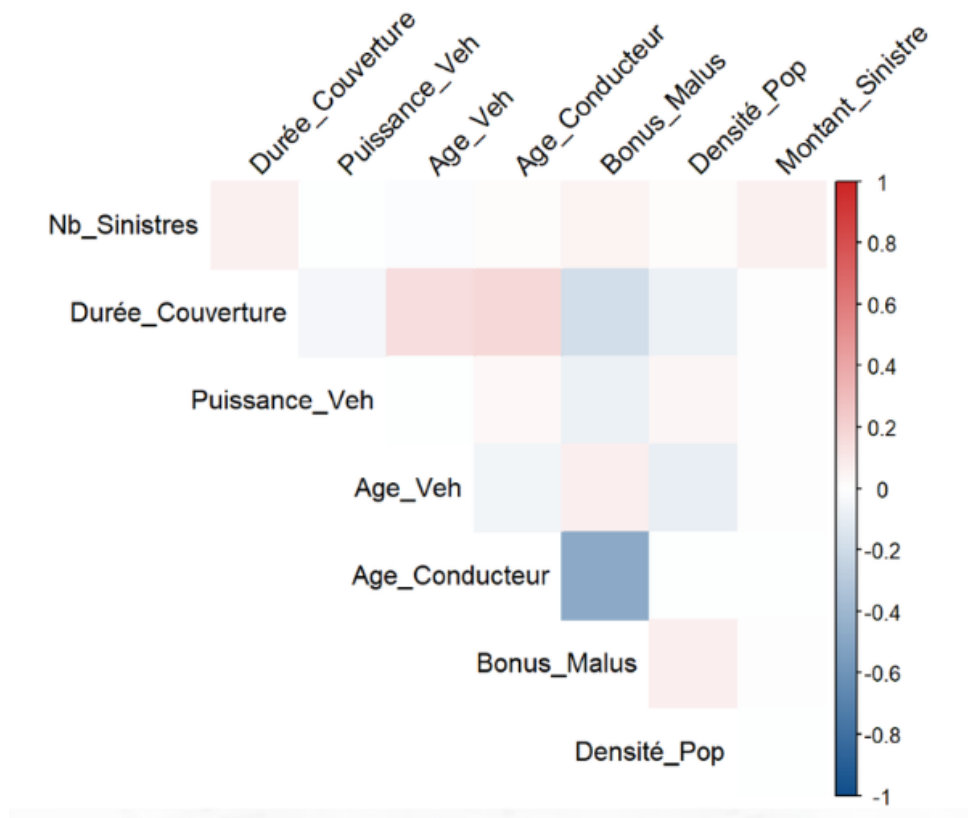
On observe également que les véhicules assurés sont majoritairement récents et possèdent le plus souvent une puissance moyenne.

La grande majorité des contrats sont des contrats annuels, ce qui est classique en assurance auto.

De plus, la quasi-totalité des assurés dispose d'un bonus de 50, le bonus maximum, ce qui indique qu'il s'agit de conducteurs plutôt prudents.

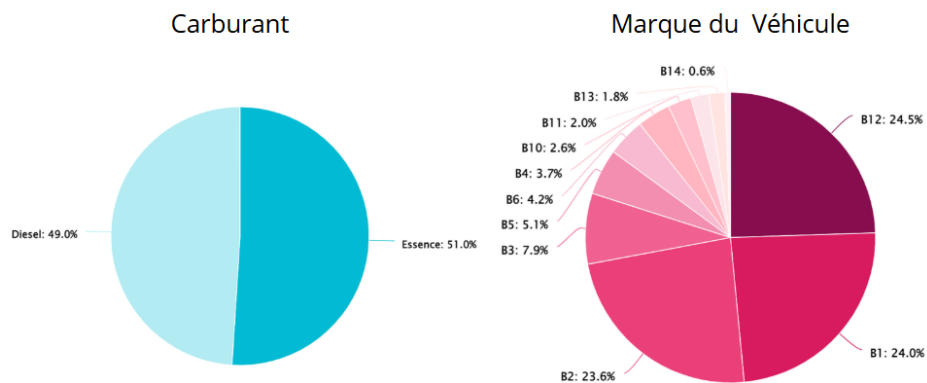
Enfin, on voit que les assurés vivant dans des zones très denses sont minoritaires.

Matrice de corrélation



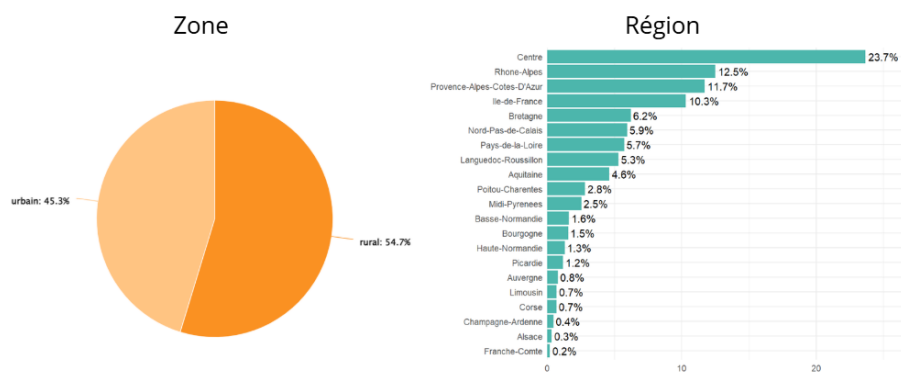
Pour aller plus loin, nous allons maintenant examiner quelles variables sont réellement liées entre elles. On constate qu'il n'existe pas de corrélation forte entre les variables. Les couleurs étant très pâles, cela indique que, dans l'ensemble, les variables sont relativement indépendantes les unes des autres. On observe seulement une légère corrélation négative entre l'âge du conducteur et le Bonus/Malus, ce qui suggère que plus les conducteurs sont âgés, plus leur bonus est élevé. Cela est logique, car les conducteurs expérimentés ont généralement moins de sinistres.

3.2.2 Variables qualitatives



On voit que pour le type de carburant, la répartition est quasiment équilibrée entre essence et diesel, avec un très léger avantage pour l'essence.

Concernant les marques de véhicules, trois catégories dominent largement : B1, B2 et B12, qui regroupent à elles seules près des trois quarts des véhicules, tandis que les autres catégories sont beaucoup moins représentées. Comme il s'agit de codes et non de vraies marques de voitures, cette variable n'apporte pas d'information pertinente et ne sera donc pas utilisée dans la suite de l'analyse.



On observe que la majorité des assurés vivent en zone rurale, un peu plus de 54

Enfin, une grande partie des assurés résident en Région Centre. Cette variable ne sera pas non plus retenue dans l'analyse, car la variable "Zone" fournit déjà le même type d'information.

Chapitre 4

Résultats des modèles

4.1 Résultats et Sélection du Modèle de Fréquence

Nous avons mis en concurrence trois modèles de comptage pour la variable dépendante Nb_Sinistres : la Loi de Poisson (1), la Loi Binomiale Négative (2) et le Modèle Hurdle (3).

	<i>Dependent variable:</i>		
	Nb_Sinistres		
	<i>Poisson</i>	<i>negative binomial</i>	<i>hurdle</i>
	(1)	(2)	(3)
Puissance_Veh	0.012*** (0.003)	0.012*** (0.003)	-0.096*** (0.010)
Age_Veh	-0.021*** (0.001)	-0.022*** (0.001)	0.014*** (0.003)
Age_Conducteur	0.014*** (0.0005)	0.014*** (0.0005)	0.002* (0.001)
Bonus_Malus	0.024*** (0.0004)	0.024*** (0.0004)	0.007*** (0.001)
CarburantEssence	-0.053*** (0.013)	-0.053*** (0.014)	0.116*** (0.037)
Zoneurbain	0.115*** (0.014)	0.117*** (0.015)	0.756*** (0.040)
Densité_Pop	-0.003 (0.017)	-0.00001 (0.019)	-0.181*** (0.048)
Constant	-4.660*** (0.044)	-4.688*** (0.049)	-16.377 (54.996)
Observations	679,513	679,513	679,513
Log Likelihood	-92,753.100	-91,699.400	-151,079.100
theta		0.582*** (0.019)	
Akaike Inf. Crit.	185,522.200	183,414.800	
<i>Note:</i>			<i>p<0.1; p<0.05; p<0.01</i>

4.1.1 Évaluation du Modèle de Poisson

Le modèle de Poisson (1) affiche la valeur de *Log Likelihood* la plus faible ($-92\,753,100$) parmi les trois modèles considérés. Bien qu'il constitue un point de référence utile pour l'interprétation des facteurs explicatifs, sa performance relative limitée était attendue. En effet, la loi de Poisson repose sur l'hypothèse d'équidispersion ($E[N] = \text{Var}[N]$). Or, dans un portefeuille d'assurance, on observe dans la majorité des cas, une forte hétérogénéité des risques conduisant à une surdispersion ($\text{Var}[N] > E[N]$). Cette caractéristique est confirmée empiriquement par l'estimation du paramètre de dispersion θ (theta) dans le modèle Binomial Négatif (2), lequel est statistiquement significatif, mettant ainsi en évidence la présence de surdispersion dans les données. Par conséquent, l'hypothèse d'équidispersion sous-jacente au modèle de Poisson apparaît inadaptée, ce qui entraîne une sous-estimation de la variance et de l'incertitude. Le modèle de Poisson est donc écarté au profit de modèles plus flexibles.

4.1.2 Comparaison Binomial Négatif versus Hurdle

Nous comparons les modèles capables de gérer la surdispersion : le Binomial Négatif (2) et le modèle Hurdle (3).

- **Critère du Log Likelihood (Log Vraisemblance)** : Le Log Likelihood mesure la qualité de l'ajustement du modèle aux données. Plus sa valeur est proche de zéro, meilleur est l'ajustement.
 - Binomial Négatif (2) : Log Likelihood = $-91\,699,400$
 - Modèle Hurdle (3) : Log Likelihood = $-151\,079,100$

Le Modèle **Binomial Négatif** présente un Log Likelihood significativement plus élevé (moins négatif), indiquant qu'il est beaucoup mieux ajusté aux données que le Modèle Hurdle.

- **Critère d'Akaike (AIC)** : L'AIC (Akaike Information Criterion) pénalise les modèles trop complexes. Le modèle avec l'AIC le plus faible est le meilleur.
 - Binomial Négatif (2) : AIC = $183\,414,800$
 - Modèle de Poisson (1) : AIC = $185\,522,200$

Bien que l'AIC du modèle Hurdle ne soit pas directement affiché, le Log Likelihood beaucoup plus faible que celui du modèle Binomial Négatif garantit que son AIC serait également très supérieur. Le Modèle Binomial Négatif est clairement le meilleur des trois en termes de parcimonie et d'ajustement.

4.1.3 Conclusion sur le Choix du Modèle

Compte tenu de ses excellentes performances en termes de Log Likelihood et d'AIC, et de sa capacité à gérer la surdispersion (θ), le Modèle Binomial Négatif est retenu pour modéliser la Fréquence des Sinistres (N_t).

4.2 Interprétation du Modèle de Fréquence

Suite à la phase de sélection du modèle (vs Poisson et Hurdle), nous comparons maintenant la performance du Modèle Binomial Négatif standard (colonne 1) à celle du Modèle Binomial Négatif Zero-Inflated (colonne 2) pour s'assurer du choix final.

TABLE 4.1 – Comparaison Binomial Négatif vs Zero-Inflated

	<i>Dependent variable:</i>	
	Nb_Sinistres	
	<i>negative</i>	<i>zero-inflated</i>
	<i>binomial</i>	<i>count data</i>
	(1)	(2)
Puissance_Veh	0.012*** (0.003)	-0.036*** (0.005)
Age_Veh	-0.022*** (0.001)	-0.002 (0.003)
Age_Conducteur	0.014*** (0.0005)	-0.0001 (0.001)
Bonus_Malus	0.024*** (0.0004)	0.012*** (0.001)
CarburantEssence	-0.053*** (0.014)	0.215*** (0.023)
Zoneurbain	0.117*** (0.015)	0.251*** (0.030)
Densité_Pop	-0.00001 (0.019)	-0.092*** (0.023)
Constant	-4.688*** (0.049)	-3.761*** (0.097)
Observations	679,513	679,513
Log Likelihood	-91,699.400	-148,761.800
theta	0.582*** (0.019)	
Akaike Inf. Crit.	183,414.800	
<i>Note:</i>	<i>p</i> <0.1; <i>p</i> <0.05; <i>p</i> <0.01	

4.2.1 Conclusion Définitive sur le Choix du Modèle de Fréquence

Le choix final du modèle se base sur l'AIC (Akaike Information Criterion), qui doit être minimisé :

- Modèle Binomial Négatif standard (1) : AIC = 183 414, 800
- Modèle Binomial Négatif Zero-Inflated (2) : Log Likelihood = -148 761, 800

Le Log Likelihood beaucoup plus faible du modèle Zero-Inflated $(-148\,761,800)$ comparé au modèle Binomial Négatif standard $(-91\,699,400)$ indique un ajustement beaucoup moins bon malgré sa complexité accrue. Par conséquent, le Modèle Binomial Négatif standard est définitivement retenu comme modèle de comptage final.

4.2.2 Synthèse des Coefficients du Modèle de Fréquence Retenu

L'interprétation se concentre sur les coefficients du Modèle Binomial Négatif (colonne 1), qui expliquent l'espérance de la fréquence $E[N_{\text{sinistres}}]$:

$$\log(E[N_{\text{sinistres}}|\mathbf{X}]) = \alpha + \log(\text{Durée_Couverture}) + \sum_i \beta_i \cdot X_i$$

1. **Caractéristiques du Véhicule** : Le coefficient de la Puissance_Veh ($\beta_1 = 0,012$) est positif, indiquant qu'une augmentation de la puissance du véhicule est associée en moyenne à une fréquence de sinistres plus élevée. À l'inverse, l'Age_Veh ($\beta_2 = -0,022$) est négatif : les véhicules plus anciens présentent donc une fréquence de sinistres légèrement plus faible, ce qui peut refléter une utilisation plus prudente ou moins intensive.
2. **Profil du Conducteur** : L'Age_Conducteur ($\beta_3 = 0,014$) est positif, suggérant que les conducteurs plus âgés ont, en moyenne, un nombre attendu de sinistres légèrement supérieur. Le coefficient du Bonus_Malus ($\beta_4 = 0,024$) est également positif, confirmant qu'un historique défavorable augmente le risque attendu et agit comme un indicateur synthétique du comportement du conducteur.
3. **Facteurs Environnementaux** : Le coefficient de Zoneurbain ($\beta_6 = 0,117$) est positif, ce qui indique que vivre en zone urbaine augmente la fréquence des accidents, probablement en raison de la densité du trafic. Le CarburantEssence ($\beta_5 = -0,053$) est associé à une fréquence de sinistres plus faible par rapport aux véhicules diesel.

4.3 Modélisation de la Sévérité des Sinistres (X_i)

La Sévérité, X_i , est le coût unitaire moyen des sinistres. Elle est modélisée par un Modèle Linéaire Généralisé (GLM) avec une loi de distribution adaptée (ici, Log-normale) et une fonction de lien logarithmique. La variable dépendante est $\log(\text{Montant_Sinistre})$.

	<i>Dependent variable:</i>
	$\log(\text{Montant_Sinistre})$
Puissance_Veh	0.009*** (0.003)
Age_Veh	-0.009*** (0.001)
Age_Conducteur	0.006*** (0.001)
Bonus_Malus	0.006*** (0.0004)
CarburantEssence	-0.014 (0.014)
Zoneurbain	0.017 (0.016)
Densité_Pop	-0.047** (0.019)
Constant	6.242*** (0.050)
Observations	26,444
R ²	0.010
Adjusted R ²	0.010
Residual Std. Error	1.125 (df = 26436)
F Statistic	37.779*** (df = 7; 26436)
Note:	$p < 0.1$; $p < 0.05$; $p < 0.01$

La structure du modèle est la suivante :

$$\log(E[X_i|\mathbf{X}]) = \gamma_0 + \sum_{i=1}^k \gamma_i \cdot X_i$$

4.3.1 Interprétation des Coefficients du Modèle de Sévérité

L'interprétation se base sur les coefficients (colonne unique) :

1. **Variables Significatives ($p < 0.01$)** : Les facteurs suivants ont un impact statistiquement démontré sur le coût moyen du sinistre :
 - Puissance_Veh ($\gamma_1 = 0,009$) : Un coefficient positif. Les véhicules plus puissants sont associés à des coûts de sinistres plus élevés.
 - Age_Veh ($\gamma_2 = -0,009$) : Un coefficient négatif. Les véhicules plus âgés sont associés à des coûts de sinistres plus faibles.
 - Age_Conducteur ($\gamma_3 = 0,006$) : Un coefficient positif. L'augmentation de l'âge du conducteur est associée à des coûts plus élevés.
 - Bonus_Malus ($\gamma_4 = 0,006$) : Un historique défavorable (Malus) est associé à des coûts unitaires plus élevés.
 - Densité_Pop ($\gamma_5 = -0,047$) : Un coefficient négatif. Plus la densité de population est élevée, plus le coût moyen du sinistre a tendance à être plus faible. Ceci peut refléter le fait que les accidents en zone très dense surviennent à faible vitesse.
2. **Variables Non Significatives** :
 - CarburantEssence et Zoneurbain : Ces facteurs, qui étaient significatifs pour la Fréquence, ne montrent pas d'impact statistique sur le coût moyen (Sévérité) du sinistre.

4.3.2 Limitation du Modèle : Le R^2 faible

Le Modèle de Sévérité présente un coefficient de détermination R^2 de seulement 0,010, soit 1%.

- **Interprétation** : Cette valeur extrêmement faible indique que les caractéristiques individuelles et géographiques de l'assuré (les \mathbf{X} que nous avons utilisées) n'expliquent que 1% de la variabilité du coût du sinistre.
- **Conclusion Actuarielle** : Ce résultat est très courant en modélisation de la Sévérité. La gravité (le coût) d'un accident dépend avant tout de facteurs stochastiques et aléatoires (angle de l'impact, présence de piétons, conditions météorologiques au moment de l'accident, etc.) qui ne sont pas capturés dans nos données de souscription. Même si un modèle futur parvenait à doubler ce R^2 à 2%, la majeure partie de la variance resterait inexpliquée.

Conclusion

Dans ce travail, nous avons présenté la théorie moderne du portefeuille développée par Harry Markowitz comme cadre conceptuel de référence pour l'analyse de la gestion de portefeuille. Fondée sur le compromis rendement–risque et sur le principe de diversification, cette théorie constitue un socle fondamental pour comprendre les mécanismes d'optimisation des décisions d'allocation des ressources face à l'incertitude.

L'analyse a montré que, bien que la théorie de Markowitz ait été initialement conçue pour les marchés financiers, ses fondements peuvent être mobilisés de manière pertinente dans le cadre de la gestion d'un portefeuille d'assurance automobile. En effet, l'assureur est confronté simultanément à des risques financiers liés à la gestion de ses placements et à des risques techniques découlant de la fréquence et de la sévérité des sinistres automobiles. L'intégration de ces dimensions permet d'élargir la notion de risque au-delà de la seule volatilité financière et de mieux appréhender les interactions entre les différentes sources d'incertitude.

Toutefois, les limites du modèle classique, notamment l'hypothèse de normalité des rendements et la prise en compte exclusive du risque financier, rendent nécessaire le recours à des approches plus avancées intégrant des risques multiples. L'utilisation de données réelles d'assurance automobile et de modèles de comptage pour la fréquence des sinistres, couplés à des modèles de sévérité, met en évidence l'intérêt d'une gestion de portefeuille adaptée aux spécificités du secteur assurantiel, où la recherche de performance doit impérativement être conciliée avec les exigences de solvabilité et de stabilité financière.

Bibliographie

Mangram, M. E. (2013). *A simplified perspective of the Markowitz portfolio theory*. Global Journal of Business Research, 7(1), 65–76.

Cobbaut, R., Gillet, R., & Hübner, G., avec la collaboration de van den Berg, A. (2è éd.). (n.d.). *La gestion de portefeuille : Instruments, stratégie et performance*. PORTEF. ISBN 978-2-8041-9012-5.

Coulon, J. (2009). *Volatilité et gestion de portefeuille* (Mémoire de doctorat, Université Claude Bernard–Lyon 1, n° 068-2009).

Poncet, P., & Portait, R. (2009). *La théorie moderne du portefeuille : Théorie et applications* (2è éd., p. 795). Groupe Eyrolles.

Elton, E. J., & Gruber, M. J. (1997). *Modern portfolio theory, 1950 to date*. Journal of Banking & Finance, 21, 1743–1759.

Hurson, C., & Zopounidis, C. (n.d.). *Gestion de portefeuille et analyse multicritère*. FeniXX.