

***Évaluation de l'actif des sociétés d'assurance
dans le contexte réglementaire
et calibrage des modèles***

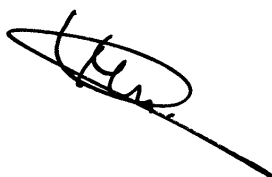
Mémoire d'actuariat présenté pour l'obtention du :

***Master science de gestion, finance de marché,
spécialité « Actuariat » du CNAM***

et l'admission à ***l'Institut des Actuaire***s

Mémoire soutenu le 3 avril 2012

Par Laëtitia VIEN :



Tuteurs : Jean-Luc BLACHON
Véronique MORAND

Jury

Président : Michel FROMENTEAU

Membres : Alexis DUPONT
Pierre PETAUTON
Vincent RUOL
François WEISS



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

Autorisation de publication et de mise en ligne sur un site de diffusion de documents actuariels

***« Remember that all models are wrong ;
the practical question is how wrong do they have to be to not be useful. »***

*George E. P. Box, Norman R. Draper
Empirical Model-Building and Response Surface, Wiley, 1987.*

Résumé

Alors que la crise financière se fait tenace, de nombreuses réformes réglementaires interviennent, notamment dans le domaine de l'assurance. Les assureurs sont détenteurs de nombreux titres d'actifs visant à faire face aux engagements envers les assurés. Les actifs s'affichent donc dans le bilan des assureurs, pour cela, ils doivent être évalués. La tendance des nouvelles réglementations est d'imposer la valeur de marché comme référence.

Outre la valeur à aujourd'hui du titre, l'assureur cherche à projeter à court, moyen ou long terme la valeur de son portefeuille afin de mieux appréhender les risques qu'il supporte et d'optimiser sa structure.

L'étude va s'attacher à évaluer la valeur d'un portefeuille fictif à l'aide de modèles stochastique, ceux-ci étant capables de simuler des grandeurs économiques en introduisant un aléa.

Nous verrons comment choisir un modèle, tester sa qualité et enfin l'impact que peut avoir de mauvais choix.

Abstract

While the financial crisis is tough, many regulatory reforms take place, in particular in the field of insurance. Insurance companies are holding many titles of assets to be able to meet commitments to policyholders. Therefore, assets appear on the balance sheet of insurers, for this they must be evaluated. The trend of the new regulations is to require the market value as the reference.

In addition to the present value of the title, the insurer seeks to project the short, medium or long term value of its portfolio to better understand the risks that it supports.

The study will focus on evaluating the value of a hypothetical portfolio using stochastic models, they are able to simulate economic variables by introducing vagaries.

We will see how to choose a model, test its quality and finally the potential impact of bad choice.

Remerciements

Je tiens à remercier en quelques mots toutes les personnes qui ont contribués de près ou de loin à l'élaboration de ce document.

Je tiens à exprimer ma profonde gratitude à mes directeurs de mémoire :

- ✓ Jean-Luc BLACHON (statisticien), pour tous ces conseils avisés, son soutien technique et sa disponibilité ;
- ✓ Véronique MORAND (actuaire) pour ces conseils et sa relecture attentive.

Mes remerciements s'adressent aussi à Hervé MORAND, Michel BOITEUX, et Véronique MORAND associés du groupe Sinalys, qui m'ont permis de pouvoir suivre les enseignements dispensés au CNAM durant deux années.

Enfin, je remercie mes collègues et mon entourage pour les encouragements et le soutien sans faille qu'ils m'ont donnés pour aboutir dans ce projet.

Sommaire

I.	Introduction	1
A.	Contexte réglementaire.....	1
1.	ACP – États règlementaires.....	2
2.	IASB – IFRS	3
3.	EIOPA – Solvabilité II	4
4.	CFO Forum – Embedded Value	7
B.	Problématique de l'étude et enjeux	8
1.	Problématique	8
2.	Enjeux	10
3.	Déroulement	11
II.	Les générateurs de scénarios économique.....	12
A.	Définition d'un GSE.....	12
B.	Générateur de Wilkie.....	15
1.	Présentation du modèle	15
2.	Détail des équations	16
3.	Bilan sur le modèle de Wilkie.....	21
C.	Générateur Ahlgrim et Al.....	23
1.	Présentation du modèle	23
2.	Détail des équations	24
3.	Bilan sur le modèle d'Ahlgrim et Al.....	27
III.	Méthode de calibrage des modèles d'actifs	28
A.	Estimation des paramètres du modèle.....	28
1.	Méthode d'estimation	28
2.	Matrice de variance – covariance / corrélation	33
3.	Ecart-type des estimateurs	34
B.	Qualité de l'ajustement du processus	34
1.	Analyse de la variance (ANOVA)	35
2.	Test Fisher Snedecor (inclus dans ANOVA)	36
3.	Coefficients de déterminations (R^2 et $R^2_{ajusté}$)	37
4.	Test de Student.....	37
C.	Test des hypothèses du modèle	38
1.	Test de normalité des résidus.....	38
2.	Test de normalité standardisé des résidus	41
3.	Test d'indépendance des résidus (autocorrélation)	42
4.	Test sur l'hypothèse des résidus identiquement distribués	44
D.	Particularité d'un modèle intégrant des sauts : Merton	45
1.	Présentation du modèle	46
2.	Estimation par la méthode des moments.....	48
3.	Estimation par le maximum de vraisemblance	49
4.	Qualité d'ajustement du modèle et vérification des hypothèses	50
E.	Critères de comparaison des modèles.....	54
1.	Critère Likelihood Ratio Test (LRT).....	55
2.	Critère Akaike (AIC / AICc)	55
3.	Critère Bayésien (BIC)	56
IV.	Mise en œuvre du calibrage des classes d'actifs	57
A.	L'inflation	58
1.	Les données	58
2.	Modèle de Vasicek.....	58
3.	Résultat du calibrage	66
B.	Les taux d'intérêts (court et long)	70
1.	Les données	70
2.	Modèle de Vasicek à deux facteurs	71
3.	Résultat du calibrage	78

C.	Les actions	82
1.	Les données	83
2.	Modèle de Black & Scholes	84
3.	Modèle de Merton	84
4.	Résultat du calibrage – série de Friggitt	84
5.	Résultat du calibrage – série d’AXA	87
D.	L’immobilier	89
1.	Les données	89
2.	Modèle de Vasicek	91
3.	Modèle de Black & Scholes	91
4.	Résultat du calibrage	91
V.	Impact du calibrage sur un portefeuille	95
A.	Le portefeuille et éléments de simulation	95
1.	Composition du portefeuille	95
2.	GSE retenu	96
3.	Matrice de corrélation	96
4.	Construction des variables aléatoires	97
B.	Comparaison de deux calibrages	97
1.	Le rendement de l’inflation	98
2.	Le rendement des taux (court et long)	99
3.	Le prix des obligations	100
4.	Le rendement des actions Friggitt	101
5.	Le rendement de l’action AXA	102
6.	Le rendement de l’immobilier	103
C.	Impact du calibrage sur le portefeuille global	103
VI.	Conclusion.....	106
VII.	Annexes.....	107
A.	Modèles AR / MA / ARCH	107
1.	Modèle Auto Régressif d’ordre p (ARp)	107
2.	Modèle Moyenne Mobile d’ordre p (MAp)	107
3.	Modèle AutoRegressive Conditional Heteroskedasticity d’ordre p (ARCHp)	107
B.	Algorithme et tests	108
1.	Algorithme de GAUSS - NEWTON	108
2.	Principe des tests d’hypothèse	109
3.	Test Fisher Snedecor (inclu dans ANOVA).....	109
C.	Données	110
1.	Données de Friggitt	110
2.	Données de l’action AXA	111
D.	Détail des calculs d’estimation	113
1.	Modèle de Vasicek à un facteur.....	113
2.	Modèle de Vasicek à deux facteurs	127
3.	Modèle de Merton	128
VIII.	Références.....	138
A.	Sites	138
B.	Ouvrages.....	138
C.	Documents.....	138
D.	Mémoires.....	140

I. Introduction

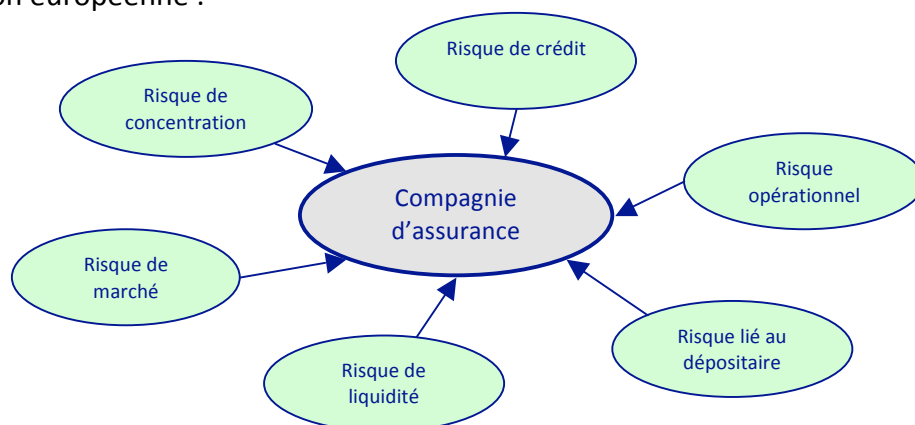
L'enjeu d'un assureur ou réassureur est de savoir évaluer les risques inhérents à son activité afin de pouvoir allouer un capital suffisant pour les couvrir, tout en conservant une rentabilité suffisante. L'objectif du mémoire est de proposer une méthode d'évaluation des actifs qui puisse être utilisée pour l'évaluation de ce capital. Il est aussi de prendre en compte les contraintes réglementaires, qui vont fortement évoluer avec la mise en place de la directive Solvabilité II, afin de permettre une évaluation des risques associés par une calibration des plus adaptées.

A. Contexte réglementaire

Les sociétés d'assurance sont particulièrement suivies et contrôlées par l'Etat du fait de l'impact que pourrait avoir la faillite d'une société sur ses assurés et sur l'économie en général. La principale difficulté, pour suivre cette activité, est le cycle de production inversé (les sinistres liés à un contrat d'assurance ne sont pas connus au moment de la tarification du contrat), c'est pourquoi il est nécessaire d'identifier les risques et de les analyser le plus précisément possible.

La crise de 2008-2009 a mis en lumière l'incidence sur les compagnies d'assurance que pouvaient avoir des chocs sur le marché financier. Lorsque l'évaluation des actifs est faite en valeur de marché, le résultat des compagnies est alors beaucoup plus volatile. Plus récemment, les difficultés rencontrées par Groupama sont liées, notamment, à la composition de son actif, à la fois dans l'exposition aux dettes souveraines, à une participation importante dans des sociétés dont la cotation en bourse a été fortement malmenée et des investissements déficitaires à l'étranger.

Voici une représentation des risques majeurs observés pendant et suite à la crise, selon la commission européenne :



Un des risques majeurs est le risque de marché avec l'évaluation des actifs que possède une société d'assurance. Ce risque est lui-même découpé en sous risques (variation du taux d'intérêt, du taux de change, des actions, etc...).

Le mémoire traitera plus particulièrement l'évaluation des actifs financiers. Le code des assurances indique à ce sujet que « *Les entreprises d'assurance doivent procéder en permanence à une évaluation de leurs risques financiers en effectuant notamment des simulations de l'impact de la variation des taux d'intérêt et des cours boursiers sur leur actif et leur passif et des estimations comparées de l'exigibilité de leur passif et de la liquidité de leur actif.* » R332-1-2, l'article R332-10-1 concernant les réassureurs.

La même exigence est imposée dans le Code de la Mutualité par les articles R212-30 pour les assureurs et R212-59-1 pour les réassureurs. Dans le code de la Sécurité Sociale, les articles correspondants sont R931-47-10-1 pour les réassureurs et R931-47-10-2 pour les assureurs.

De nombreux états de contrôle sont produits pour déterminer la robustesse d'une compagnie, chacun poursuivant un objectif différent :

- Les comptes sociaux qui sont contrôlés et validés par les Commissaires aux comptes ;
- Les autres états réglementaires qui permettent à l'autorité de contrôle de vérifier la capacité des compagnies d'assurance à respecter leurs engagements vis-à-vis des assurés ;
- Lorsqu'ils sont établis, les comptes en normes IFRS permettent d'avoir une vision proche de la valeur de marché pour la communication au marché financier.

Et bientôt :

- Les états du pilier 3 de la directive Solvabilité 2;
- La généralisation de l'Embedded Value et de la MCEV (Market Consistent Embedded Value) qui permettra de comparer les valeurs de portefeuilles et de sociétés d'assurance entre elles.

1. ACP – États réglementaires

En France, l'organisme spécifique au contrôle des sociétés d'assurance est l'**ACP** (Autorité de Contrôle Prudentiel), qui est notamment responsable du contrôle des sociétés d'assurance, des mutuelles et des institutions de prévoyance.

L'évolution du contexte économique ces dernières décennies tend à la création de sociétés liant des activités de banque et d'assurance. Ce constat a naturellement mis en évidence qu'il devrait exister un lien plus fort entre les contrôles purement assuranciers et ceux qui relèvent des activités bancaires ou financières. A ce titre, il a été créé l'Autorité de contrôle prudentiel (**ACP**) née de la fusion des autorités d'agrément et de contrôle de la banque (CECEI Comité des établissements de crédit et des entreprises d'investissement) et de l'assurance (ACAM / CEA Comité des entreprises d'assurance). La publication au Journal officiel date du 22 janvier 2010.

Afin de prouver sa capacité à tenir ses engagements envers les assurés, une société doit fournir une série d'états réglementaires de périodicité variable (trimestrielle ou annuelle). Les différents Codes régissant les activités d'assurance (Code des Assurances, Code de la Sécurité Sociale, Code de la Mutualité), définissent le contenu de ces états.

Trois états sont en lien direct avec l'évaluation de l'actif d'une société d'assurance :

- **Etat C5** : Représentation des engagements règlementés – Cet état permet d'évaluer la couverture des engagements de l'assureur, en confrontant les passifs règlementés avec les actifs en représentation des engagements.
- **Etat T3** : Simulations actif-passif - Les entreprises établissent trimestriellement, un état retraçant l'incidence de chocs sur la valeur de réalisation de leurs placements ainsi que sur leurs provisions mathématiques, suivant les hypothèses imposées par la réglementation.
- **C6 bis** : Stress test - Il sert à mesurer la capacité des assureurs à honorer leurs engagements à 5 ans dans des conditions de sinistralité et d'environnement financier conjointement dégradées. La grandeur qui ressort au final est le montant de l'éventuelle moins-value qui résulterait de la cession immédiate de la partie des actifs nécessaire à payer les prestations durant ces 5 années.

Ils complètent le bilan et l'état détaillé des placements et son tableau complémentaire, le TCEP, qui décrivent le contenu du portefeuille.

Par ailleurs, une baisse importante de la valeur des actifs peut avoir un impact sur le bilan des assureurs pour l'évaluation des provisions, dont :

- **PAF** (Provision pour aléa financier) : provision constituée lorsqu'il existe un risque d'écart entre les taux d'intérêt technique et les taux garantis.
- **PDD** (Provision pour dépréciation durable) : provision constituée lorsque la valeur d'une action non stratégique demeure inférieure à sa valeur au bilan de manière durable.
- **PRE** (Provision pour risque d'exigibilité) : provision constituée lorsque la valeur globale des placements est inférieure à la valeur au bilan.
- **PGG** (Provision globale de gestion) : cette provision est destinée à couvrir les charges de gestion futures des contrats qui ne seraient pas couvertes par ailleurs (selon l'article R 331-3 du Code des assurances), notamment les frais de gestion des placements.

D'autres demandes ou enquêtes peuvent concerner les actifs. Certaines débouchent sur l'ajout de nouveaux états comme celle diligentée pour le 15 juin 2010 concernant l'état détaillé des placements des sociétés d'assurance (TCEP).

2. IASB – IFRS

Les entreprises françaises d'assurance peuvent être soumises à deux types de normes comptables : la norme sociale française et les normes internationales.

Les normes de comptabilisation internationales sont définies par l'IASB (*International Accounting Standard Board, anciennement IASC*), organisme qui définit les règles comptables IAS (*International Accounting Standards*) devenues les normes IFRS

(*International Financial Reporting Standard*). Leur mise en place doit faciliter la comparaison internationale et harmoniser les pratiques.

Depuis 2005, les sociétés européennes cotées ou faisant appel à de l'argent public, doivent publier leurs comptes dans ce référentiel, l'information délivrée est principalement destinée aux marchés financiers. Les éléments comptables sont alors évalués en valeur de marché (principe de *fair value*).

Les sociétés d'assurance, de par leur cycle de production inversé, ont des normes plus particulièrement établies à leur intention comme l'IFRS 4, qui concerne l'évaluation de leur passif.

D'une manière générale, la liste ci-après peut impacter directement l'évaluation des compagnies :

- **IAS1** : pour la présentation des états financiers
- **IAS 19** : pour la comptabilisation des avantages sociaux¹
- **IAS 32** : présentation des instruments financiers
- **IAS 39** : pour la valorisation des instruments financiers

D'autres sont en cours de discussion :

- **IFRS 4** : pour la valorisation des passifs d'assurance
- **IFRS 7** : Information à fournir sur les instruments financiers

Les normes de comptabilisation poursuivent leur évolution. A ce titre l'IASB a publié, le 12 novembre 2009, un article visant à remplacer l'IAS 39 par l'IFRS 9. La refonte de cette norme devrait permettre de réduire la complexité des évaluations actuelles et ainsi améliorer la lisibilité des valeurs des instruments financiers.

Les normes **IFRS4 phase 2** et **IFRS 9** devaient entrer en vigueur au 1^{er} janvier 2013, mais ces dates sont encore incertaines. En particulier, l'IASB a publié en août 2011 un exposé sondage pour le report de l'application des IFRS 9 aux comptes de 2015.

En s'intéressant de plus près à l'IFRS 9, il ressort que l'évaluation des actifs financiers serait faite de deux façons :

- En coût amorti (prix d'achat + frais d'acquisition – éventuelle dépréciation) : pour les instruments ayant les caractéristiques de base d'un prêt et gérés sur la base de flux contractuels, par exemple les obligations cotées indexées sur l'inflation.
- A la juste valeur (cours du titre ou prix auquel il serait négocié, ceci peut recouvrir la valeur de marché ou la valeur de réalisation,...) : pour les autres instruments, par exemple les obligations cotées indexées sur les taux swaps.

Cette évaluation en juste valeur tranche avec les normes de comptabilisation française, où le coût d'achat de l'actif était pris en compte (valeur historique) en opposition à la valeur de marché (valeur aujourd'hui). De nouvelles méthodes de comptabilisation des actifs doivent être mises en place au sein des compagnies d'assurance pour répondre à ce besoin.

3. EIOPA – Solvabilité II

Solvabilité I

¹ Avantages sociaux correspond ici à toutes formes de contrepartie données par l'entreprise en échange d'un service rendu par les employés.

Depuis les années 1973 (pour la non-vie) et 1979 (pour la vie), les sociétés d'assurance se doivent de constituer un niveau minimum de fonds propres afin de couvrir les engagements pris envers les assurés. Ce capital est appelé marge de solvabilité, il doit être représenté par des actifs adéquats, notamment dans la période résiduelle de couverture de l'engagement.

En France, cette marge est calculée proportionnellement aux provisions mathématiques en assurance vie ou des primes et des sinistres en non-vie.

Afin de prouver que les contraintes réglementaires sont respectées, un rapport doit être remis chaque année aux autorités concernées.

Les critères de l'exigence de capital du régime actuel de solvabilité s'exposent à plusieurs critiques dont les principales sont :

- Certains risques sont pas suffisamment pris en compte (ex : chocs sur les marchés)
- Il n'existe pas d'homogénéité de calculs d'un pays à l'autre

Solvabilité II

Suite aux critiques du modèle de Solvabilité I, la commission européenne a délégué au CEIOPS (*Committee of European Insurance and Occupational Pensions Supervisors*) la réforme de la directive Solvabilité I. Ce comité rassemblant les autorités de contrôle de l'Union Européenne, doit promouvoir la directive Solvabilité II qui devra être mise en application début 2013.

Ce comité est devenu l'EIOPA (*European Insurance and Occupational Pensions Authority*) au 1^{er} janvier 2011, suite à la réforme des structures de supervisions financières européennes.

Les objectifs de la directivité Solvabilité II sont principalement :

- Consolider un marché commun de l'assurance, pour favoriser le développement commercial entre les pays (ex : LPS), tout en intégrant les particularités de chacun d'entre eux (ex : le régime de sécurité sociale en France). De plus la protection de l'assuré s'en trouvera renforcée face à la plus grande mobilité des citoyens.
- Replacer les risques au cœur du système des règles prudentielles.
- Harmoniser les règles d'évaluation et de contrôles des sociétés au sein de l'Union pour créer un marché Européen.

L'homogénéisation des règles d'évaluation et de contrôles adaptées aux risques des assureurs, s'articulent autour de trois piliers dans la directive :

- Pilier 1 : Exigences en capital, il s'agit de la définition d'une exigence quantitative de couverture de la solvabilité avec le MCR (Minimum Capital Requirement) et le SCR (Solvency Capital Requirement)
- Pilier 2 : Exigences qualitatives, c'est-à-dire un renforcement du contrôle interne et de la gouvernance, un contrôle accru de la qualité des données, une mise en place de modèles permettant de mieux appréhender le risque à gérer, une pratique de l'auto-évaluation des besoins en capital, etc...
- Pilier 3 : Information prudentielle et informations à destination du public (autorité de contrôle, communication financière)

La directive a été adoptée, mais la finalisation de ses mesures d'application n'est pas achevée, elle se poursuit conformément au processus Lamfalussy². Cette réforme se définit au fil d'échanges entre l'EIOPA et le marché de l'assurance (sociétés, fédérations d'assureurs, organisations professionnelles...). Les échanges sont effectués par l'intermédiaire d'études quantitatives d'impact (QIS), commencées en 2005, ainsi que de consultations du marché, les *consultation papers*. Les assureurs ont répondu en 2010 au 5^{ème} QIS, chaque échange permet d'améliorer le modèle théorique proposé afin de mieux traduire les réalités des sociétés et de converger vers une formule d'appréciation du niveau de solvabilité cohérente et partagée.

L'EIOPA travaille notamment à la définition des exigences en capital : le MCR représentant le niveau minimum de fonds propres en dessous duquel l'intervention de l'autorité de contrôle sera automatique pour un retrait d'agrément (15% de probabilité de ruine à 1 an), et le SCR représentant le capital cible nécessaire pour absorber un choc exceptionnel (0,5% de probabilité de ruine à 1 an). Comme l'énonce le troisième pilier, les sociétés assurances seront tenues de communiquer leur situation financière, à l'ACP et au public, avec de plus en plus de transparence et de clarté, et de prouver leur robustesse. La communication d'éléments financiers au public sera beaucoup plus large qu'elle ne l'est aujourd'hui, à la fois dans son contenu et par le nombre de sociétés concernées : certaines d'entre elles, comme les mutuelles relevant du Code de la Mutualité, ne diffusent pas leurs comptes actuellement.

Le calcul du SCR est défini par modules de risques, celui qui va nous intéresser dans le cadre de cette étude est le module du risque de marché. Ce module représente quantitativement une composante importante du SCR, que ce soit en assurance vie ou en assurance non-vie.

Il est découpé en sous modules :

- Risque de taux d'intérêt : mesure l'impact d'une variation de la courbe des taux d'intérêt sur la valeur de l'actif net de l'entreprise.
- Risque action : lorsque l'actif net est sensible à la volatilité du cours des actions.
- Risque immobilier : lorsque l'actif net est sensible à la volatilité du cours de l'immobilier.
- Risque de devise : lorsque l'actif net est sensible à la volatilité du cours des taux de change.
- Risque de contrepartie : mesure les pertes qui pourraient être engendrées par le défaut ou la dégradation de la notation de la contrepartie.
- Risque de concentration des marchés : risque lié à l'accumulation des expositions sur une même contrepartie ou sur des contreparties liées, appelé aussi risque idiosyncratique qui résulte donc d'une diversification insuffisante.

Concernant la valorisation des actifs et passifs, l'article 75 de la directive Solvabilité II définit les principes généraux à suivre. Les actifs doivent être valorisés au montant auquel ils pourraient être échangés dans le cadre d'une transaction conclue entre les parties. Les passifs sont valorisés au montant auquel ils pourraient être transférés ou réglés dans le

² **Processus Lamfalussy** : permet de mettre en place des textes en quatre étapes :

- 1) Directives ou règlements (grand principe)
- 2) Mesures d'exécution Directives ou règlements (plus détaillé que dans la première partie)
- 3) Recommandations non contraignantes
- 4) Contrôles renforcés des infractions potentielles

cadre d'une transaction conclue entre les parties. C'est donc une valorisation de marché qui est demandée.

En réalité, pour la directive Solvabilité II, on peut retenir deux modes d'évaluation des actifs :

- « mark to market » : lorsque la valeur des actifs est observable sur le marché, ou lorsqu'il en existe possédant les mêmes caractéristiques
- « mark to model » : lorsque le prix n'est pas directement observable ou qu'il n'existe pas d'actif similaire, l'actif sera évalué par actualisation des flux à l'aide d'un modèle mathématique.

4. CFO Forum – Embedded Value

La notion d'Embedded Value, notée EV, a été introduite pour mesurer la valeur des actifs nets des sociétés d'assurance du point de vue des actionnaires, dans les années 90. Il s'agit de la valeur intrinsèque des sociétés, sans goodwill.

En effet, la valeur d'une société d'assurance ne peut être calculée que sur le long terme du fait de ses engagements envers les assurés. Il n'est pas aisé pour des financiers d'appréhender tous les risques inhérents aux engagements d'une société d'assurance, c'est pourquoi l'EV s'est petit à petit imposée comme l'indicateur incontournable de communication à destination des analystes financiers et du marché. Cependant, les méthodes mises en place à l'époque par les compagnies n'étaient pas homogènes.

En 2002, un groupe de discussion européen des assurances a été créé, le **CFO Forum**, dans lequel on retrouve les directeurs financiers de grandes sociétés d'assurance européennes cotées. Son but est d'améliorer la cohérence, la transparence et la comparabilité de l'information financière des assureurs, afin de mettre en confiance les investisseurs qui pouvaient avoir des réticences face à cette difficulté à appréhender les risques.

Dans les faits, la mise en place de l'Embedded value a suscité et suscite encore beaucoup d'interrogations quant à la meilleure façon de la calculer. C'est pourquoi le CFO Forum, a émis différentes versions de l'EV à mesure que les réflexions ont avancé :

- TEV : Traditionnal Embedded Value
- EEV : European Embedded Value
- MCEV : Market Consistent Embedded Value

Market Consistent Embedded Value - MCEV

Pour résoudre les problèmes de comparabilité des sociétés d'assurance, le CFO Forum a introduit la notion de MCEV (Market Consistent Embedded Value) afin d'évaluer le portefeuille des compagnies en valeur de marché, dite « Market Consistent ».

La MCEV représente alors la « valeur des intérêts des actionnaires dans les revenus distribuables issus des actifs alloués au business couvert » [*principe 3*].

La première étape est alors de mettre en place un modèle actif/passif comme l'ALM (Asset Liability Management), pour les compagnies n'en possédant pas, pour avoir une bonne vision des risques de leur portefeuille. L'ALM permettra de mieux piloter l'activité et les risques de la société par :

- Une meilleure adaptation de l'allocation stratégie des actifs (adossement actif/passif)

- La définition des orientations commerciales vers lesquelles il faudra tendre
- Une communication financière de qualité

Ces méthodes favorisent l'utilisation des modèles stochastiques notamment pour évaluer les actifs en valeur de marché et avoir une vision prospective du portefeuille.

Concrètement, les différentes publications du CFO Forum concernant la modélisation des actifs restent silencieuses sur les modélisations à utiliser. La modélisation doit permettre l'évaluation « Market consistent » et donc intégrer la possibilité de choc sur le marché c'est-à-dire utiliser la modélisation stochastique, comme le confirme le *Principe 15* « Les modèles stochastiques doivent être calibrés à partir des données de marché observables et les plus récentes possibles. Il en va de même des hypothèses sur les volatilités implicites, elles doivent être déduites des prix des instruments dérivés échangés sur le marché plutôt que sur les observations historiques des taux d'intérêts ou des actifs sous-jacents. »

Finalement, tous les flux futurs doivent être évalués avec la valeur d'actifs similaires sur le marché. Les hypothèses de projection peuvent s'appuyer sur des données internes ou externes à la compagnie, mais doivent être appropriés à l'activité couverte. Les hypothèses doivent être revues régulièrement.

B. Problématique de l'étude et enjeux

1. Problématique

Comme nous venons de l'exposer dans le contexte réglementaire, les nombreuses normes régissant le monde de l'assurance convergent vers une demande d'évaluation des actifs en valeur de marché.

Dans ce contexte, il devient incontournable pour les sociétés de se doter d'outils performants permettant d'évaluer les risques au plus juste, notamment pour les actifs qu'ils détiennent.

Il ressort de ces premières recherches, qu'il existe un certain vide dans les normes quand à l'évaluation des actifs, ce qui laisse une grande liberté au niveau du choix d'évaluation de chaque classe d'actifs, tant sur les méthodes que sur les modèles.

Il est à noter que certaines contraintes devront être respectées pour se conformer à la réglementation, comme :

- Ne pas sous estimer les pertes afin d'obtenir une bonne évaluation de la probabilité de ruine à 99.5% pour le SCR ;
- Disposer de visions historiques des actifs ;
- Intégrer les contraintes de projections, court terme et long terme (jusqu'à 30 ou 40 ans) ;
- Inclure l'inflation qui peut avoir un impact tant sur les primes que les prestations, et les dépendances entre actifs.

Nous allons donc chercher comment répondre au mieux à la demande d'évaluation des actifs en valeur de marché.

Pour être en mesure d'estimer la valeur future d'un portefeuille d'actif, il est nécessaire de simuler des variables dont la fluctuation des valeurs impacte directement la valeur des actifs (par exemple : l'inflation, le taux d'intérêt...). Ces projections doivent pouvoir tenir compte d'hypothèses économiques si nécessaire, telles que la prise en compte de périodes de crise.

L'utilisation des générateurs de scénarios économiques (GSE) s'impose alors comme une évidence puisqu'ils permettent de projeter des grandeurs économiques à plus ou moins long terme à l'aide de modèles dans le but d'étudier l'évolution de la valeur des actifs financiers.

Dans la littérature, plusieurs GSE ont été développés, chacun proposant une structure ou des modèles différents, il nous appartiendra de choisir celui qui nous semble le plus approprié et/ou de proposer des alternatives si les modèles ne nous semblent pas adaptés aux hypothèses économiques.

Chacun des risques identifiés doit pouvoir être modélisé par une ou plusieurs variables incluses dans le(s) modèle(s).

On doit tenir compte des influences que les variables économiques simulées peuvent avoir entre elles (par exemple, lorsque le taux d'inflation augmente, le taux d'intérêt augmente bien souvent aussi). Nous retiendrons l'hypothèse selon laquelle l'influence entre les variables est linéaire, elle pourra être matérialisée par une matrice de corrélation. L'étude de Kamal Armel (juin 2010) montre que cette hypothèse n'est pas totalement réaliste en raison de la mauvaise prise en compte des dépendances de queues de distribution, donc des périodes de crises, l'alternative qu'il propose est l'utilisation des copules qui pourrait être une piste d'approfondissement de cette étude.

Ces générateurs étant constitués de modèles, il faudra ensuite les calibrer. Et c'est bien là tout l'enjeu de notre étude.

Il existe deux types de calibrage, le calibrage statistique et le calibrage de marché.

A priori, le calibrage de marché n'est pas très adapté à l'estimation d'un portefeuille d'assurance puisqu'il se base uniquement sur les valeurs des actifs sur le marché à un instant t pour déterminer les paramètres du modèle financier.

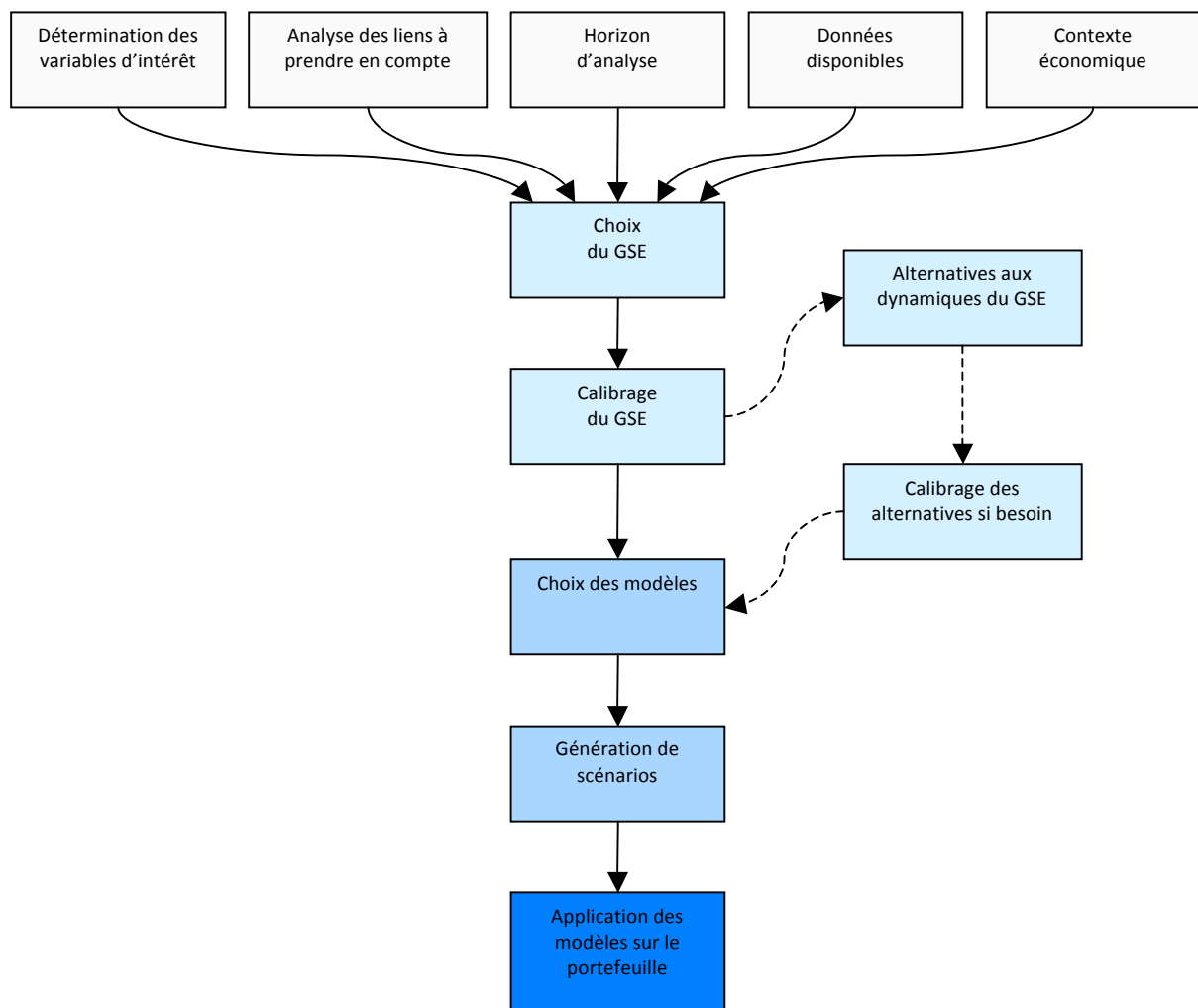
Un portefeuille d'assurance n'étant pas destiné à être vendu dans l'instant, il est préférable de s'appuyer sur un historique de données dont la profondeur reste à déterminer. C'est donc naturellement que la calibration statistique va être étudiée.

Concrètement il s'agit d'obtenir l'évaluation des paramètres du modèle mais aussi de disposer des éléments permettant de déterminer la fiabilité du modèle obtenu.

Pour cela nous allons proposer une méthode de calibrage qui se veut la plus complète possible, en plusieurs phases :

1. L'évaluation des paramètres
2. L'étude de la qualité d'ajustement du modèle aux données
3. La vérification de la validité des hypothèses du modèle
4. La comparaison des modèles si nécessaire

La démarche peut être schématisée :



2. Enjeux

L'enjeu est important, un mauvais choix de modèle ou une mauvaise évaluation des paramètres des modèles impliquerait des simulations d'actifs qui ne reflèteraient pas les hypothèses économiques posées.

De ce fait, les différents calculs prudentiels se basant sur les résultats de simulations ne seront pas significatifs, voir complètement aberrants.

Les conséquences sont d'autant plus visibles en période de crise. Avec la crise actuelle, beaucoup de modèles financiers ont été remis en cause car il s'est avéré qu'aucun n'avait prévu un tel « chaos ». De fait, les sociétés d'assurance n'avaient pas préparé de modèles résistant à un tel choc.

Par exemple, les sociétés d'assurance utilisent bien souvent le modèle Black & Scholes pour procéder à l'évaluation des actifs. Ce modèle s'appuie sur l'hypothèse que les rendements des actifs suivent une loi log-normale, empiriquement cette hypothèse n'est pas toujours vérifiée. Les queues de distributions sont souvent plus épaisses, ceci signifie que le modèle ne prend pas suffisamment en compte la possibilité de choc sur le marché.

A ce propos, le CEIOPS a demandé qu'en décembre 2009, un stress-test soit mené sur les plus grands et plus importants groupes d'assurances européens, afin d'évaluer l'impact d'une grande crise sur la solvabilité des assureurs selon trois scénarios différents :

- Un choc financier du type de celui constaté entre septembre 2008 et septembre 2009
- Une récession plus profonde et prolongée
- Un scénario de reprise de l'inflation et des taux d'intérêts

L'objectif du calibrage est de pouvoir choisir le modèle le mieux adapté aux attentes, aux données, aux risques. Il est important de garder en mémoire le fait que tous les modèles sont faux mais qu'il faut déterminer le moins mauvais. Dans les statistiques, il faut aussi prévoir ce qui n'est jamais arrivé, donc quelle que soit la qualité de l'évaluation d'actifs obtenus, il faut rester vigilant et interpréter les résultats avec suffisamment de recul.

3. Déroulement

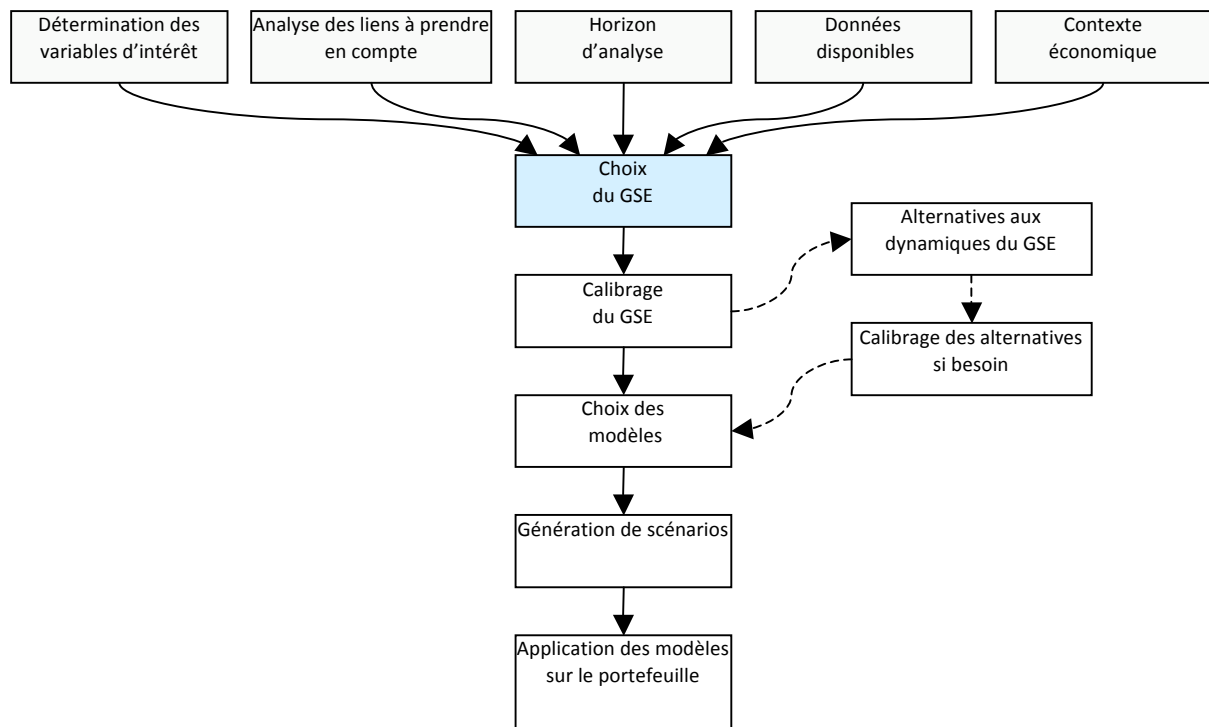
La première étape consiste à étudier des GSE. Nous avons retenu deux des plus connus, à savoir celui de Wilkie et celui d'Ahlgrim.

La partie suivante sera consacrée à la présentation de la méthode de calibrage, ce qui inclut l'estimation des paramètres, la qualité d'ajustement et la vérification des hypothèses.

Nous poursuivrons par une mise en application des méthodes de calibrage sur les différentes classes d'actif à partir des données de Friggitt et de l'action AXA. Enfin nous conclurons par l'impact de mauvais choix de modèles, ou de calibrage sur l'évaluation d'un portefeuille.

II. Les générateurs de scénarios économique

L'utilisation des scénarios économiques est une solution pour répondre aux exigences d'évaluation en valeur de marché. Il convient maintenant de définir précisément ce qu'est un GSE, et de déterminer celui qui nous semble le plus adapté à notre problématique.



A. Définition d'un GSE

Un générateur de scénarios économique (GSE) permet de prendre en compte un ou plusieurs facteurs (taux d'inflation, taux d'intérêt, etc...) qu'on nommera macro variable, afin de faire des projections de leurs valeurs à plus ou moins long terme, dans le but d'étudier l'évolution de la valeur d'un actif financier par exemple.

Un GSE est constitué de plusieurs équations différentielles, chacune exprimant la dynamique d'une macro-variable. C'est par ce biais qu'on peut introduire l'aléa ou les chocs possibles sur les valeurs d'actifs dans le futur.

Le modèle choisi doit être une représentation adéquate du passé, tout en étant capable de réaliser des projections tenant compte d'hypothèses économiques. En outre, pour avoir une vision fiable d'un résultat futur (ou d'une estimation de la Value-At-Risk (VaR)), il est nécessaire de procéder à de nombreuses simulations.

La détermination du nombre de simulation suffisant, permettant d'avoir un résultat cohérent, est un compromis entre le nombre minimum nécessaire et la limitation des moyens techniques (le temps de calcul peut être démultiplié, il faut donc limiter le nombre de simulations).

Le choix d'un générateur de scénarios économiques n'est pas chose aisée, il suppose de résoudre plusieurs questions :

- 1) Quelle est la vocation du générateur ?

- 2) Quelles variables macro-économiques ou financières souhaite-t-on modéliser (taux d'intérêt, taux d'inflation, rendements actions, etc...)?
- 3) Quel(s) modèle(s) doit-on choisir pour la dynamique d'évolution des variables précitées ?
- 4) Quelle structure de projection des variables doit-on choisir ?
- 5) Considère-t-on que les différentes variables économiques sont liées et de quelle manière peut-on les modéliser ?
- 6) Souhaite-t-on évaluer des actifs à court terme ou long terme ?
- 7) Quelles données a-t-on à notre disposition ?

1. Vocation du générateur

Le choix des composantes d'un GSE est supposé être lié à sa vocation finale que ce soit en tant qu'outil d'évaluation des produits financiers pour les stratégies d'achat ou de vente (pricing), ou en tant qu'outil de projection et de gestion des risques.

Nous cherchons ici à trouver un bon générateur pour le cas de la projection des valeurs d'actifs.

2. Macro variables pouvant être intégrées au modèle

Les macro-variables qu'il est pertinent d'intégrer dans le modèle sont celles qui ont une influence sur la valeur finale d'un actif. Si on recense *a priori* les facteurs les plus influents dans ce domaine, on a :

- Le taux d'inflation,
- Le taux d'intérêt nominal à court ou long terme,
- Les rendements des actions,
- Les rendements de l'immobilier
- La croissance économique,
- Le taux de change,
- Le taux d'inflation des salaires,
- Etc...

C'est à l'aide des simulations de certaines de ces macro-variables qu'il sera possible d'évaluer la valeur future des actifs et passifs d'une société d'assurance.

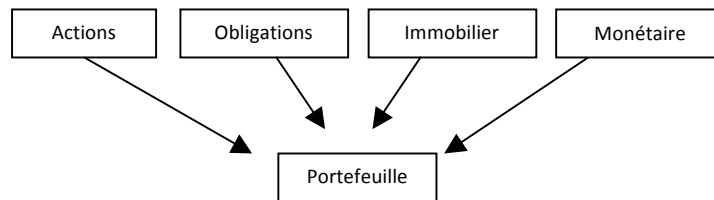
Il est important de pouvoir tenir compte dans les modèles d'autres facteurs comme un changement de politique gouvernementale ou une innovation technologique, qui peuvent avoir une forte influence sur les valeurs économiques.

3. Structure du modèle

Chacune des variables va être modélisée au travers d'équations autorégressives, différentielles, etc... Lors de cette étape, il est important de tenir compte de la dépendance des différentes variables économiques entre elles. La structure du GSE permet d'en tenir compte, il en existe principalement deux types :

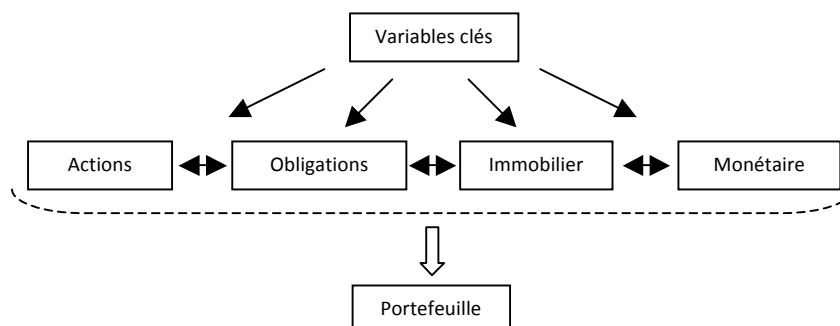
- Modèle intégré : structure en cascade dans laquelle une ou plusieurs variables « motrices » vont permettre de déduire les valeurs des autres variables. Par exemple, partant du taux d'inflation on déduit les taux d'intérêts, les taux de

dividendes, etc...c'est le cas du modèle de Wilkie (1986), ou de Brennan et Xia (2000). Ce type de GSE a l'avantage d'avoir une implémentation simple, avec une compréhension des interactions entre variables lisibles. En revanche, ils possèdent beaucoup de limites, entre autres, en termes de capacité prédictive.



Remarque : Modélisation séparées de chaque catégorie d'actifs

- **Modèle composite :** structure par dépendance reposant sur l'idée que les variables puissent avoir des dépendances de manière simultanée (ex : l'évolution du rendement action et de l'inflation peuvent être liées). Dans ce type de modèle, on spécifie chaque classe d'actif avant de modéliser les structures de dépendances (ex : corrélations, copules) entre elles. C'est le cas du modèle d'Ahlgrim et Al. (2005). Ce type de modèle possède lui aussi des inconvénients comme sa difficulté de mise en œuvre et une forte dépendance aux données sur lesquelles il est basé.



Remarque : Modélisation des variables clés, puis modélisations des catégories d'actifs en tenant compte des interactions.

4. Structure de projection des variables

Après avoir choisi la structure du GSE, décrite précédemment, il faut déterminer une structure de projection pour chaque variable. Bien souvent, on retient une même structure de projections pour toutes les variables.

Selon le travail présenté par Planchet, Faleh, Rullière (2009), il en existe deux : structure linéaire et structure en arbre.

Chaque date observée est représentée par un nœud, à partir de chaque nœud il peut exister une ou plusieurs trajectoires possibles suivant la structure choisie :

- **Structure linéaire des scénarios :** une seule trajectoire est dérivée à partir de chaque nœud.
- **Structure d'arbre des scénarios :** chaque nœud possède différents nœuds enfants, donc plusieurs trajectoires possibles.

La seconde proposition est bien plus complexe, elle peut aussi apporter des problèmes techniques car le nombre de scénarios est rapidement démultiplié, le temps de calcul peut devenir explosif. Dans les deux GSE étudiés, la structure linéaire est retenue.

5. Evaluation d'actif à court, moyen ou long terme ?

Un GSE de bonne qualité doit pouvoir réaliser des projections fiables sur un moyen/long terme, tout en intégrant les conséquences que pourraient avoir de grosses variations sur le court terme.

Pour les besoins d'évaluation technique en assurance, il est important d'avoir une bonne vision de la valeur de l'entreprise à moyen terme (15 ans), voir long terme. En effet, certaines provisions que l'assureur doit constituer sont à long terme, comme par exemple pour les contrats de type : retraite, épargne, assurance décès.

6. Données

Lorsqu'on a étudié de manière théorique les moyens mathématiques nous permettant de réaliser des simulations pour les macro-variables souhaitées, il faut prendre en ligne de compte les données.

Le choix des outils mathématiques peut être influencé par la disponibilité ou la qualité des données. Des problématiques peuvent se poser quant à la profondeur d'historique (travailler sur le long terme exige de posséder une longue période d'observation), actifs ayant une liquidité insuffisante, données confidentielles pour certains types d'actif, etc...

La seconde chose à vérifier, ce sont les hypothèses du modèle étudié, il ne faut pas quelles soient trop restrictives ou invérifiables dans la pratique. A titre d'exemple, on peut évoquer que des théories reposent souvent sur la normalité des données, dans les faits, l'hypothèse est bien souvent non vérifiée, il faut donc en tenir compte dans le modèle choisi.

Dans la suite, nous allons présenter plus particulièrement deux générateurs de scénarios, le générateur de Wilkie (1986 / 1995), puis celui de Ahlgrim et Al. (2004), en essayant d'en exposer les avantages et inconvénients afin de savoir lequel des GSE sera le plus à même de répondre à notre problématique.

B. Générateur de Wilkie

1. Présentation du modèle

Le générateur, développé par A. D. Wilkie, publié en 1986, est principalement basé sur l'indice des prix à la consommation (IPC), dont on déduit l'inflation. Cette variable est importante car elle peut impacter autant la valeur des primes que des prestations chez un assureur, ainsi que la valeur des actifs.

Ce GSE est un modèle en cascade, le principe est le suivant : à partir des données d'historique, on calibre le modèle de l'IPC pour en obtenir les paramètres. Puis avec le modèle ainsi calibré, on pourra effectuer les projections pour l'IPC et en déduire toutes les autres projections de variables. Donc toutes les projections de variables de ce modèle découlent de l'IPC.

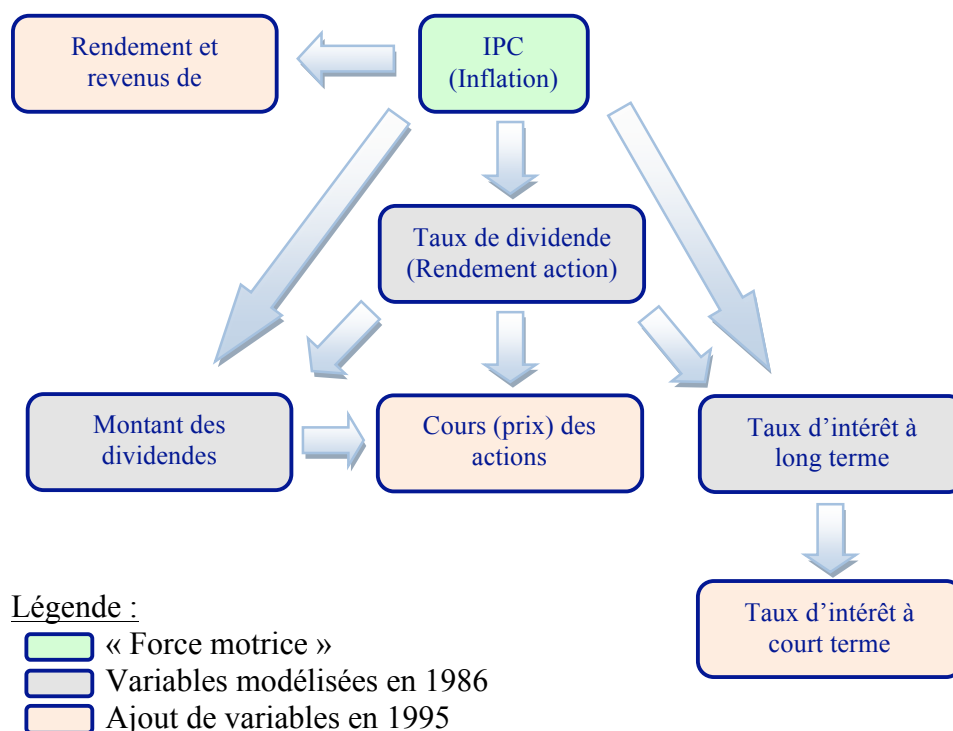
A sa parution, ce modèle était novateur car il était l'un des premiers à intégrer l'ensemble des variables nécessaires pour décrire l'investissement dans un fond à savoir :

- L'IPC (inflation),
- L'indice des dividendes,
- Rendement des actions,
- Taux d'intérêt (taux moyen d'emprunt d'État).

En 1995, Wilkie retravaille son modèle pour l'enrichir des variables suivantes :

- Cours des actions,
- Taux d'intérêt à cours terme,
- Rendement de l'immobilier.

Structure du GSE Wilkie (1986 - 1995)



2. Détail des équations

a) Modélisation du taux d'inflation

Pour modéliser l'inflation, Wilkie propose une équation basée uniquement sur l'historique des données (moyenne + un pourcentage de déviation de l'année précédente) additionné à un bruit blanc (constituant l'introduction d'aléa, de moyenne nulle).

L'équation du taux d'inflation est définie par une équation autorégressive :

$$I(t) = \mu_I + \alpha_Q [I(t-1) - \mu_I] + \sigma_Q * \varepsilon_Q(t)$$

Avec :

μ_I taux d'inflation long terme,

α_Q déviation de l'IPC par rapport à l'année antérieure, l'introduction d'un facteur de retour à la moyenne s'explique par le fait que des valeurs élevées de taux sont souvent suivies par des baisses (et inversement), de sorte que les taux restent toujours autour d'une moyenne.

σ_Q écart type de la variable $\varepsilon_Q(t)$,

$\varepsilon_Q(t) \sim N(0,1)$ bruit blanc gaussien, indépendamment et identiquement distribuée (*iid*)

Les paramètres à estimer sont : $\mu_I, \alpha_Q, \sigma_Q$.

L'équation pour la modélisation de l'IPC se présente alors sous la forme suivante :

$$Q(t) = Q(t-1) * e^{I(t)}$$

Soit encore : $I(t) = \ln Q(t) - \ln Q(t-1)$, le taux d'inflation qui est la différence de l'IPC dans l'année $(t-1, t)$.

Remarque : Il est préférable de modéliser l'inflation en premier, plutôt que l'IPC, car ce sont les variations relatives de l'IPC qu'il est important d'étudier, plutôt que les valeurs absolues qui apportent moins d'information.

Dans sa première étude, basée sur les données économiques au Royaume-Uni de 1919 à 1982, Wilkie avait modélisé l'inflation par un modèle Autorégressif d'ordre 1, noté AR(1), c'est-à-dire utilisant les données de $t-1$ uniquement (*cf. §Modèle Auto Régressif d'ordre p (ARp)*), en supposant que les résidus étaient normalement distribués.

L'amélioration du modèle en 1995, avec l'utilisation d'un modèle ARCH (AutoRegressive Conditional Heteroskedasticity) (*cf. §Modèle AutoRegressive Conditional Heteroskedasticity d'ordre p (ARCHp)*), a permis de tenir compte des caractéristiques de distributions des données d'historique. Les queues de distribution se révèlent en réalité plus épaisses que celles de la loi normale, ce constat a été prouvé dans l'étude de KITTS en 1990 « Comments on a model of retail price inflation ». Wilkie en a alors profité pour travailler sur des données d'historique et de périodicités plus importantes.

Ce modèle permet d'introduire une variance au bruit blanc, alors qu'elle est constante avec un modèle AR(1); Ainsi l'inflation peut être plus fortement impactée si une valeur « inhabituelle » a été observée l'année précédente.

b) Modélisation du prix des actions

Une fois la modélisation de l'inflation effectuée, on peut évaluer le prix des actions en trois étapes, en déterminant la dynamique :

- 1) Du rendement des actions,
- 2) Des montants relatifs de dividendes,
- 3) Du prix des actions.

(1) Dynamique du rendement action $Y(t)$

La dynamique proposée par Wilkie est composée de l'inflation actuelle, du rendement des actions de la période précédente additionné d'un bruit blanc :

$$Y(t) = \exp \left(w_Y I(t) + \underbrace{\ln \mu_Y + \alpha_Y (\ln Y(t-1) - \ln \mu_Y) + \sigma_Y \varepsilon_Y(t)}_{\text{Modèle autorégressif de la même forme que celui de l'inflation}} \right)$$

Avec $Y(t)$ rendement instantané des actions,
 w_Y sensibilité du rendement nominal des actions à l'inflation,
 μ_Y taux de rendement de long terme,
 α_Y coefficient de retour à la moyenne,
 σ_Y écart type du bruit blanc,
 $\varepsilon_Y(t) \sim N(0,1)$, variable aléatoire centrée réduite *iid* (bruit blanc).
 Les paramètres à estimer sont : $w_Y, \mu_Y, \alpha_Y, \sigma_Y$.

(2) Dynamique des montants relatifs de dividendes $K(t)$

Pour obtenir le prix des actions, il faut ensuite modéliser le montant des dividendes $D(t)$.

$$D(t) = D(t-1) * \exp[I_D(t) + \mu_D + y_D \sigma_Y \varepsilon_Y(t-1) + b_D \sigma_D \varepsilon_D(t-1) + \sigma_D \varepsilon_D(t)]$$

Avec $I_D(t) = w_D * M_D(t) + x_D * I(t)$ l'influence de l'inflation sur les dividendes en fonction des valeurs présentes et passées, sachant que $M_D = d_D * I(t) + (1 - d_D) * M_D(t-1)$. Dans la définition du modèle, Wilkie impose la contrainte : $x_D = 1 - w_D$ de sorte qu'une variation de $\ln Q(t)$ produise le même changement sur $\ln D(t)$.
 μ_D la croissance réelle moyenne du montant des dividendes,
 $y_D \sigma_Y \varepsilon_Y(t-1)$ l'innovation du rendement l'année passée
 $b_D \sigma_D \varepsilon_D(t-1)$ l'innovation de l'année t-1 pour les dividendes, on est alors dans le cas d'un processus moyenne mobile d'ordre 1, noté MA(1) (cf. §Modèle Moyenne Mobile d'ordre p (MAp)).
 $\sigma_D \varepsilon_D(t)$ l'innovation de l'année en cours (t), $\varepsilon_D(t) \sim N(0,1)$ *iid*.

Les paramètres à estimer sont : $\mu_D, y_D, b_D, \sigma_D, w_D, d_D, x_D$.

Il est préférable comme précédemment de travailler sur l'évolution relative au cours d'une année du rendement, plutôt que sur les valeurs absolues, l'équation est alors :

$$K_D(t) = \ln \left(\frac{D(t)}{D(t-1)} \right)$$

On utilise ici une transformation logarithmique afin de passer à un modèle multiplicatif qui permet d'éviter d'obtenir des prévisions négatives.

(3) Dynamique du prix des actions $P(t)$

Une fois les deux premières dynamiques déterminées, la déduction du prix des actions se fait facilement par l'équation suivante :

Rendement des actions = (Montant dividende distribué) / (Prix des actions)

Soit $Y(t) = D(t)/P(t)$, d'où :

$$P(t) = D(t) / Y(t)$$

c) Modélisation des taux d'intérêt

(1) Taux long terme $C(t)$

La dynamique des taux d'intérêt est définie en distinguant la composante de l'inflation future et celle de la part réelle du taux d'intérêt. Ceci respecte la théorie de l'« effet de Fisher », selon laquelle toute variation de l'inflation se reflète dans le taux nominal. L'équation proposée par Wilkie est la suivante :

$$C(t) = \underbrace{w_C * M_C(t)}_{\text{Part inflation}} + \underbrace{R_C(t)}_{\text{Part réelle}}$$

Avec $M_C(t) = d_C * I(t) + (1 - d_C) * M_C(t - 1)$ dynamique de la part d'inflation,

$R_C(t) = \mu_{R_C} * \exp(N_{R_C}(t))$ dynamique de la part réelle,

$$\begin{aligned} N_{R_C}(t) &= \alpha_{1_{R_C}} * (R_C(t - 1) - \mu_{R_C}) + \alpha_{2_{R_C}} * (R_C(t - 2) - \mu_{R_C}) + \alpha_{3_{R_C}} * (R_C(t - 3) - \mu_{R_C}) \\ &+ y_{R_C} * \sigma_Y * \varepsilon_Y(t) + \sigma_{R_C} * \varepsilon_{R_C}(t) \\ \varepsilon_{R_C}(t) &\sim N(0,1) \text{ iid} \end{aligned}$$

Les paramètres à estimer sont : $w_C, d_C, \mu_{R_C}, \alpha_{1_{R_C}}, \alpha_{2_{R_C}}, \alpha_{3_{R_C}}, y_{R_C}, \sigma_{R_C}$.

Le logarithme du taux réel est alors un processus AR(3) :

$$\ln(R_C(t)) = \ln(\mu_{R_C}) + N_{R_C}(t)$$

Wilkie proposera une simplification : utiliser un modèle AR(1). Concrètement cela signifie que :

$$N_{R_C}(t) = \alpha_{1_{R_C}} * (R_C(t - 1) - \mu_{R_C}) + y_{R_C} * \sigma_Y * \varepsilon_Y(t) + \sigma_{R_C} * \varepsilon_{R_C}(t)$$

Il y a dans ce cas, deux paramètres de moins à estimer : $\alpha_{2_{R_C}}, \alpha_{3_{R_C}}$.

(2) Taux long court terme $B(t)$

Wilkie a constaté ce qui est intuitif, à savoir que les taux court terme sont liés aux taux long terme, en se basant sur les séries mensuelles des taux d'intérêt d'une banque. Les taux d'intérêt court terme à la date t , sont modélisés par :

$$B(t) = C(t) * \exp(-B_D(t))$$

Avec $B_D(t) = \mu_{B_D} + \alpha_B * (B_D(t - 1) - \mu_{B_D}) + \sigma_B * \varepsilon_B(t)$ représentant la différence instantanée entre les logarithmes des taux d'intérêt à long terme $C(t)$ et les logarithmes des taux d'intérêt à court terme $B(t)$,
 $\varepsilon_B(t) \sim N(0,1) \text{ iid}$.

Les paramètres à estimer sont : $\mu_{B_D}, \alpha_B, \sigma_B$.

Remarque : B_D est négatif car les rendements court terme sont, en moyenne, plus petits que les rendements sur le long terme, ce qu'on retrouve dans la théorie économique car l'incertitude à long terme est plus importante, donc c'est plus rémunérateur pour les investisseurs.

d) Modélisation de l'immobilier

La modélisation de l'immobilier sera effectuée en trois étapes, en déterminant la dynamique :

- 1) Du rendement de l'immobilier,
- 2) Des revenus relatifs de l'immobilier,
- 3) Du prix de l'immobilier.

Wilkie fait alors le rapprochement entre la modélisation des actions et celle de l'immobilier. En effet, les loyers perçus par la location d'un bien immobilier peuvent être assimilés aux dividendes d'une action. Il en va de même pour les rendements de l'immobilier et des actions.

(1) Dynamique du rendement immobilier $Z(t)$

La dynamique proposée par Wilkie est :

$$Z(t) = \exp(w_Z I(t) + \ln \mu_Z + \alpha_Z (\ln Z(t-1) - \ln \mu_Z) + \sigma_Z \varepsilon_Z(t))$$

Avec $Z(t)$ rendement instantané de l'immobilier,

w_Z sensibilité du rendement nominal de l'immobilier à l'inflation,

μ_Z taux de rendement de l'immobilier,

α_Z coefficient de retour à la moyenne,

σ_Z écart type du bruit blanc,

$\varepsilon_Z(t) \sim N(0,1)$, variable aléatoire centrée réduite *iid* (bruit blanc).

Les paramètres à estimer sont : $w_Z, \mu_Z, \alpha_Z, \sigma_Z$.

Cette équation est similaire à celle des rendements actions $Y(t)$.

(2) Dynamique des revenus relatifs de l'immobilier $E(t)$

Pour obtenir le prix de l'immobilier, il faut ensuite en modéliser les revenus $E(t)$.

$$E(t) = E(t-1) * \exp[I_E(t) + \mu_E + \sigma_E \varepsilon_E(t)]$$

Avec $I_E(t) = w_E * M_E(t) + x_E * I(t)$, influence de l'inflation sur les dividendes en fonction des valeurs présentes et passées, sachant que $M_E = d_E * I(t) + (1 - d_E) * M_E(t-1)$.

Dans la définition du modèle, Wilkie impose, comme précédemment, la contrainte : $x_E = 1 - w_E$ de sorte qu'une variation de $\ln Q(t)$ produise le même changement sur $\ln E(t)$.

μ_E la croissance réelle moyenne du montant des revenus de l'immobilier,

σ_E l'écart type du bruit blanc,

$\varepsilon_E(t) \sim N(0,1)$ *iid*.

Les paramètres à estimer sont : $w_E, x_E, d_E, \mu_E, \sigma_E$.

Il est préférable de travailler, là aussi, sur l'évolution relative au cours d'une année du rendement, plutôt que sur les valeurs absolues, l'équation est alors :

$$K_E(t) = \ln \left(\frac{E(t)}{E(t-1)} \right)$$

(3) Dynamique du prix de l'immobilier $A(t)$

Une fois les deux premières dynamiques déterminées, la déduction du prix de l'immobilier se fait facilement par l'équation suivante :

$$A(t) = E(t) / Y(t)$$

3. Bilan sur le modèle de Wilkie

Le modèle de Wilkie est l'un des plus connu et est souvent utilisé dans les compagnies d'assurance de par les avantages décrits, ci-après. En revanche, on peut remarquer dans la littérature que ce modèle a fait l'objet de nombreuses critiques, en voici quelques unes.

a) Avantages

Les principaux avantages connus du modèle de Wilkie sont :

- Sa simplicité d'implémentation
- Sa facilité de compréhension
- Le modèle n'intégrant pas de variables exogènes, l'explication des variables du modèle ne vient pas d'une force extérieure au système considéré.

b) Inconvénients

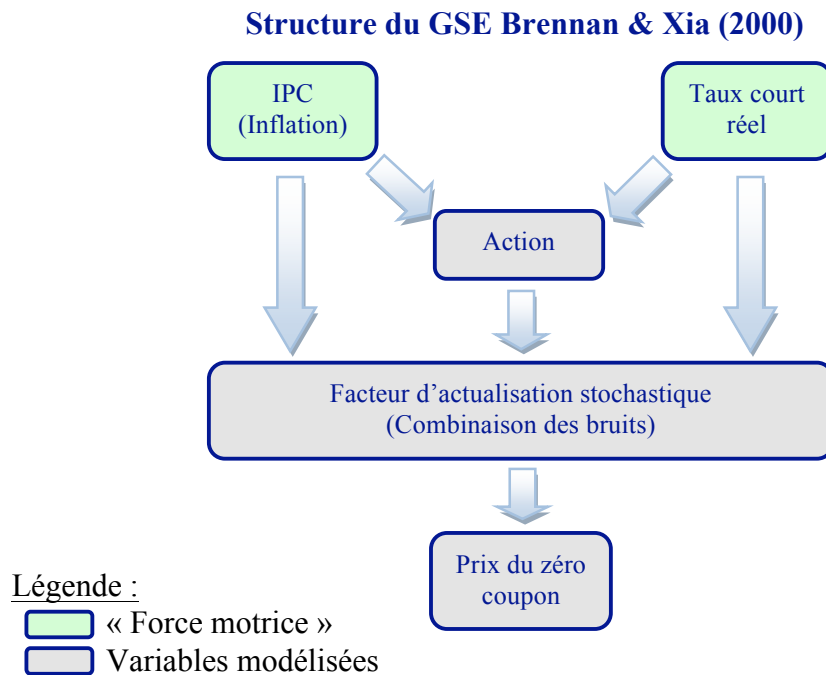
Les avantages du modèle de Wilkie sont indéniables, cependant les limitations du modèle sont quant à elles très nombreuses. A ce titre, les principaux inconvénients observés sur le modèle sont :

- De nombreux paramètres à estimer, donc la probabilité que l'un d'entre eux soit mal estimé et fausse les projections est d'autant plus élevée.
- Une faible capacité prédictive car ce modèle se base énormément sur les données observées, constat souvent exprimé, notamment par Sahin, Cairns, Kleinow, Wilkie (2008).
- L'hypothèse forte de normalité des résidus ne reflète pas la réalité, l'ensemble des estimations de paramètres et donc l'ensemble de projections s'en trouvent donc faussés. Comme l'ont démontré Sahin, Cairns et Kleinow, en 2008, en ré-estimant l'ensemble des paramètres du modèle.
- On peut constater que les projections de Wilkie, peuvent avoir assez fréquemment une inflation négative, dans les faits c'est plus rare !
- Ce GSE n'est pas adapté à l'évaluation des produits dérivés, selon Rambaruth (2003), or les sociétés d'assurances ont de plus en plus souvent recours à ce type de produits.
- La relation entre le taux d'intérêt long terme et le taux d'inflation n'est pas cohérente avec la théorie financière, Huber (1997).

c) Comparaison des caractéristiques de deux GSE intégrés

Un autre GSE intéresse les assureurs, celui de Brennan et Xia, développé en 2000, il possède beaucoup de similitudes avec celui de Wilkie.

La structure du GSE Brennan et Xia :



La comparaison ci-après s'appuie sur l'étude de Caja, Carvas et Rousseau (2009) :

Questions	GSE Wilkie	GSE Brennan Xia
Vocation	Projection de variables macro-économiques Gestion des risques	Projection de variables macro-économiques Gestion des risques
Variables « motrices »	IPC	IPC Taux court réel
Variables modélisées	IPC (inflation) Indice des dividendes Rendement des actions Taux moyen emprunt d'État Immobilier	IPC (inflation) Action Taux court réel Prix des obligations indexées sur l'inflation
Dynamique d'évolution des variables	Modèle en cascade	Modèle en cascade
Dépendance des variables	Exprimé par la corrélation entre variables	Exprimé par la corrélation entre variables
Possibilité de projection moyen ou long terme	Oui	Oui

Au vu du tableau présenté, les GSE sont proches. Ils possèdent plusieurs avantages (facilité de mise en œuvre et d'interprétation de par leur structure) et inconvénients communs (ex : l'hypothèse de résidus gaussiens).

Ils diffèrent sur d'autres points, par exemple Brennan et Xia ne modélisent pas l'immobilier, or ce type d'actif représente parfois une partie non négligeable du portefeuille d'un assureur (40% maximum).

En revanche, on peut considérer que le modèle de Brennan et Xia est plus facilement paramétrable car le nombre de paramètres à estimer est moindre, donc moins de risque d'introduction d'une erreur.

Pour poursuivre l'étude des GSE, intéressons nous maintenant à celui d'Ahlgrim et Al.

C. Générateur Ahlgrim et Al.

1. Présentation du modèle

Le modèle Ahlgrim a été développé avec le soutien de la Casualty Actuarial Society³ (CAS) et de Society of Actuaries (SOA)⁴, dans le cadre d'une étude sur les modèles DFA (Dynamic Financial Analysis) et ALM (Asset Liability Management) non-vie, c'est-à-dire la gestion actif/passif.

Ce GSE permet de remédier à certaines critiques du modèle de Wilkie :

- Relation incohérente entre le taux d'intérêt et l'inflation
- L'utilisation de modèle autorégressif qui serait trop simple au regard des données observées

Sa vocation est de réaliser des projections sur un horizon de 50 ans, en tenant compte des dépendances réelles entre les variables. L'interdépendance entre les variables se base sur la corrélation.

Les variables modélisées sont :

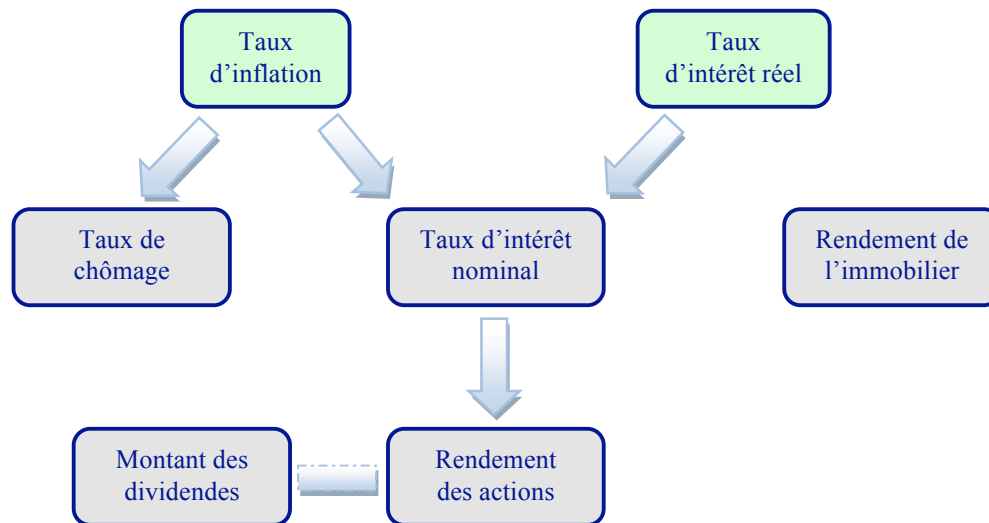
- L'inflation
- Le taux d'intérêt réel
- Le taux d'intérêt nominal
- Le rendement des actions
- Le montant des dividendes
- Le rendement de l'immobilier
- Le taux de chômage

Les relations entre ces différentes variables se présentent sous la forme:

³ Association professionnelle reconnue aux Etats Unis

⁴ Association professionnelle reconnue aux États Unis

Structure du GSE Ahlgrim et Al. (2004)



Légende :

- « Force motrice »
- Variables modélisées

Remarque : la modélisation du chômage n'est pas utile pour l'étude, elle ne sera donc pas présentée dans la suite du document.

2. Détail des équations

a) Modélisation du taux d'inflation

Dans le GSE d'Ahlgrim, l'inflation est modélisée par un processus de Vasicek à un facteur (aussi appelé Ornstein-Uhlenbeck), ce qui signifie que seul le taux court est à l'origine de la déformation de la courbe des taux.

Ce type de processus stochastiques donne lieu à des simulations Gaussiennes, c'est-à-dire que la distribution des q_t obtenue suivra une loi normale.

$$dq_t = \kappa(\mu_q - q_t)dt + \sigma dB_{q,t}$$

Avec :

κ vitesse de retour à la moyenne : des valeurs élevées de taux sont souvent suivies par des baisses (et inversement), de sorte que les taux restent toujours autour d'une moyenne.

μ_q moyenne à long terme de l'inflation autour de laquelle évolue ce taux

σ représente la volatilité

$B_{q,t}$ mouvement brownien standard permettant d'introduire l'aléa dans les simulations.

Le processus donnant des simulations gaussiennes, le taux d'inflation pourra prendre des valeurs négatives de manière relativement fréquente, ce qui ne reflète pas exactement la réalité.

La calibration du modèle peut se faire par une approximation de l'équation, proposée par Ahlgrim et Al. (2005) ou par la solution exacte présentée par Armel (2010).

b) Modélisation du taux d'intérêt réel

Le taux d'intérêt réel est modélisé par un processus de Vasicek à deux facteurs (c'est un cas particulier du modèle à deux facteurs de Hull-White).

Les deux facteurs à l'origine de la déformation de la courbe des taux d'intérêt réel sont : le taux court terme, noté r , et le taux long terme, noté l , les deux étant des variables stochastiques.

$$\begin{aligned} dr_t &= \kappa_r (l_t - r_t) dt + \sigma_r dB_{r,t} \\ dl_t &= \kappa_l (\mu_l - l_t) dt + \sigma_l dB_{l,t} \end{aligned}$$

Avec :

κ_r, κ_l vitesses de retour à la moyenne.

μ_l moyenne à long terme du taux d'intérêt réel long terme

σ_r, σ_l représentent les volatilités

$B_{r,t}, B_{l,t}$ mouvement brownien standard permettant d'introduire l'aléa dans les simulations.

La calibration de ces équations est compliquée car les données de ces variables ne sont pas directement observables sur le marché. Ahlgrim et Al. proposent alors, pour obtenir des séries de données, de partir des taux d'intérêt nominaux et d'en retrancher le taux d'inflation.

Ceci est cohérent avec la théorie économique, cette relation se nomme : l'équation de Fisher (Taux d'intérêt nominal = taux d'intérêt réel + taux d'inflation).

A l'aide des données, il est alors possible de calibrer une à une les équations car il n'est pas possible de les évaluer simultanément par la méthode des moindres carrés ordinaires (MCO) :

- Estimation des paramètres relatifs au taux d'intérêt réel long terme :

$$l_{t+1} = \beta_1 + \beta_2 l_t + \varepsilon'_{2t}$$

- Estimation des paramètres relatifs au taux d'intérêt court terme à l'aide des résultats de calibration de la première équation évaluée $\hat{l}_t$:

$$r_{t+1} = \alpha_1 (\hat{l}_t - r_t) + \varepsilon'_{1t}$$

Pour obtenir le taux d'intérêt nominal, Ahlgrim propose de combiner le taux d'inflation et le taux réel suivant la formule :

$$i = \{(1+q) * (1+r)\} - 1$$

En revanche, Armel a choisi de calculer deux taux nominaux (taux nominal court terme et long terme) en appliquant exactement la relation de Fisher décrite précédemment :

$$\begin{aligned} l_t^{no\ min\ al} &= l_t^{réel} + q_t \\ r_t^{no\ min\ al} &= r_t^{réel} + q_t \end{aligned}$$

c) Modélisation du taux de rendement des actions

(1) Le rendement hors dividendes

Ce taux est usuellement estimé en s'appuyant sur l'hypothèse de normalité des rendements actions. Des études menées sur l'historique des données ont permis de mettre en évidence que cette hypothèse n'est pas vérifiée en réalité. En effet, il a été constaté que la distribution des rendements actions avait des queues épaisses.

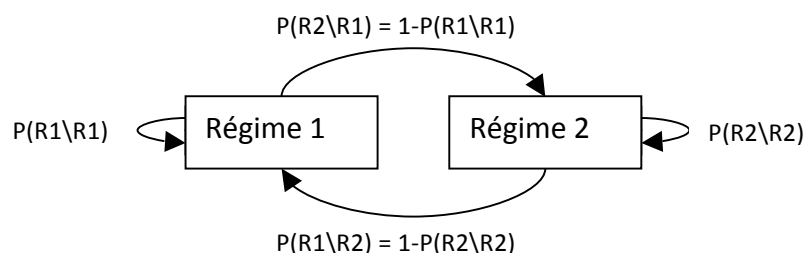
Pour modéliser le taux de rendement des actions, Ahlgrim utilise le modèle à changement de régime « regime switching », de Hardy (2001).

Il combine deux processus de Black & Scholes, l'un pour le régime « normal » avec une volatilité relativement faible et une moyenne de rendement actions relativement élevée, l'autre pour le régime de « crise » avec une forte volatilité et une moyenne de rendement actions plutôt faible. Des probabilités de transitions sont associées pour passer d'un régime à l'autre, ou rester dans l'état.

Régime 1 : $s_1 = N(\mu_1, \sigma_1)$,

Régime 2 : $s_2 = N(\mu_2, \sigma_2)$, avec $\begin{cases} \sigma_2 > \sigma_1 \\ \mu_2 < \mu_1 \end{cases}$

Schéma représentatif des probabilités de passage dans le modèle de Hardy



Le rendement hors dividendes, noté s_t , s'écrit :

$$s_t = q_t + r_t + x_t$$

Avec q_t le taux d'inflation

r_t le taux d'intérêt réel

x_t le taux de rendement en excès du taux d'intérêt réel

En outre, cette modélisation offre la possibilité de tenir compte des impacts que peut avoir la modification de la politique monétaire d'un pays, de par la possibilité de changement de régime au court du temps. Ces phénomènes ne sont pas prévisibles et constituent un aléa qui pourrait s'avérer non négligeable lorsqu'on estime la probabilité de ruine (Solvabilité 2).

(2) Le taux de dividende

La dynamique du taux de dividende action est similaire à celle présentée précédemment pour le GSE de Wilkie, bien qu'il n'intègre pas l'influence de l'inflation. Il est noté y_t , et son équation est :

$$d \ln(y_t) = \kappa_y (\mu_y - \ln(y_t)) dt + \sigma_y dB_{y,t}$$

Avec κ_y la vitesse de retour à la moyenne

μ_y le logarithme du taux de dividende moyen à long terme

Remarque : la modélisation par un modèle à deux régimes est complexe, et les données (taux de dividendes, etc...) difficiles à obtenir. De ce fait, Armel (2010), Planchet, Thérond, Kamega (2009) ont privilégié un modèle à un régime, permettant de modéliser le rendement action avec dividendes réinvestis par :

$$x_t = \ln \left(\frac{ind_t}{ind_{t-1}} \right) - r_t - q_t$$

Avec x_t la valeur historique de l'excès de rendement action,

ind_t la valeur d'un investissement en action avec dividendes réinvestis,

r_t le taux d'intérêt réel,

q_t le taux d'inflation.

d) Modélisation du taux de rendement de l'immobilier

De la même façon que pour le taux d'inflation, l'immobilier est modélisé par un processus de Vasicek à un facteur :

$$d(re_t) = \kappa_{re} (\mu_{re} - re_t) dt + \sigma_{re} dB_{re,t}$$

La tentative d'introduction d'un facteur tenant compte du taux d'inflation, dans l'étude d'Ahlgrim et Al., s'est révélée non concluante.

3. Bilan sur le modèle d'Ahlgrim et Al.

a) Avantages

Les principaux avantages dégagés par ce GSE :

- Accorder une grande importance à l'interdépendance des macro-variables, en considérant la corrélation comme instruments exprimant la liaison,
- Générateur privilégié à celui de Wilkie, qui est en plus approuvé par des associations professionnelles aux Etats-Unis, d'où une forte crédibilité.

b) Inconvénients

La principale qualité du GSE, considérer l'interdépendance des variables est aussi sa faiblesse. En effet, elle est matérialisée par la corrélation uniquement, or Armel (2010), a expliqué tous les inconvénients d'une telle modélisation. On retiendra que :

- La corrélation ne permet pas de tenir compte des dépendances des queues de distributions, donc la modélisation n'est pas pertinente en cas de crise, dans les situations extrêmes.
- La corrélation est nulle même lorsque des variables sont dépendantes (ex : $X \sim N(0,1) \Rightarrow \text{Cov}(X, X^2)=0$, pas d'indépendance X^2 n'est pas gaussienne).
- La corrélation varie par transformation logarithmique (plus généralement par transformation croissante non linéaire), ainsi $\text{corr}(X,Y) \neq \text{corr}(X)\text{corr}(Y)$, on perd donc de l'information.

III. Méthode de calibrage des modèles d'actifs

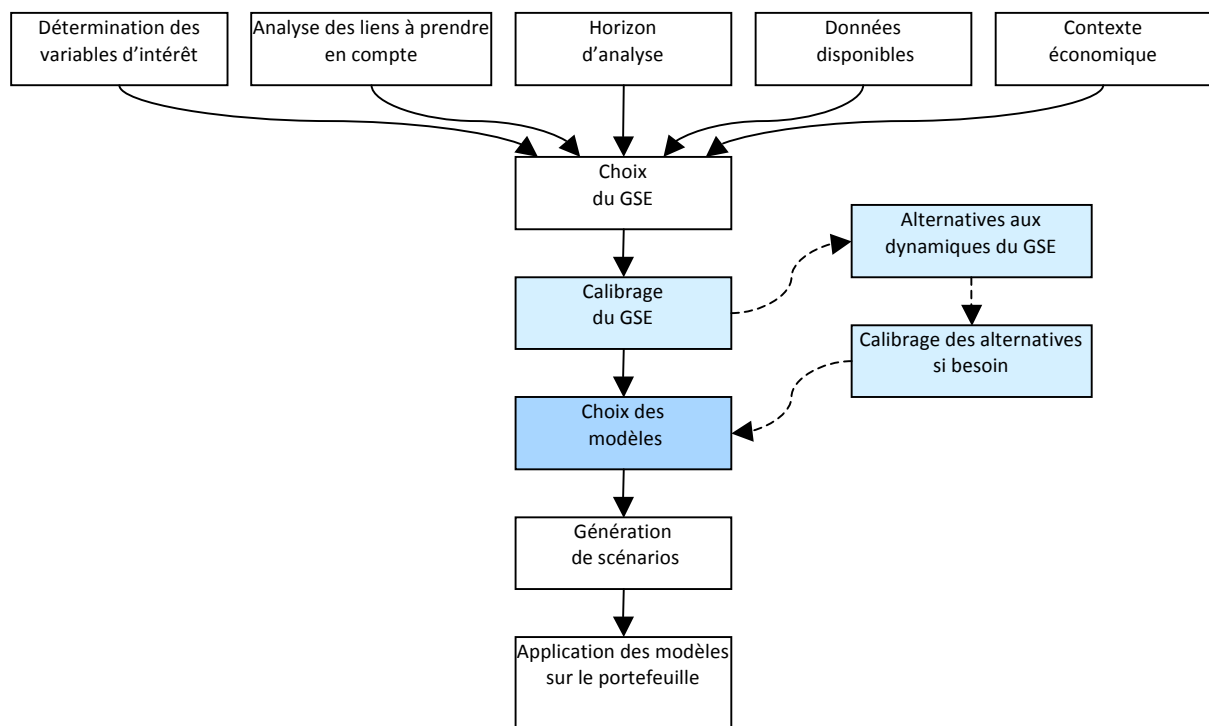
Une étape importante dans l'évaluation d'actifs est le calibrage au plus juste des modèles, qu'ils soient intégrés ou non à un GSE. La bonne réalisation d'un calibrage dépendant de plusieurs critères :

- Avoir un modèle adapté : de façon théorique, le modèle doit être cohérent avec ce qu'on souhaite évaluer
- Avoir un échantillon de données suffisant en nombre d'observation, représentatif de l'ensemble de la population
- Vérifier la validation des tests d'adéquation
- Les hypothèses du modèle doivent être testées

Le processus de calibrage va se faire en trois étapes :

- 1) L'estimation des paramètres du modèle
- 2) La qualité d'ajustement du modèle
- 3) La vérification des hypothèses du modèle

Lorsqu'une variable économique est calibrée par des modèles ou périodes de données différents, il faut être en mesure de faire un choix, nous présenterons les critères qui peuvent orienter la décision.



A. Estimation des paramètres du modèle

1. Méthode d'estimation

Dans la littérature, il existe différentes méthodes pour évaluer la valeur des paramètres d'un modèle, dont les plus connues sont :

- Les moindres carrés ordinaires, consistant à minimiser les carrés des écarts entre les valeurs observées et les estimations ;
- Le maximum de vraisemblance, consistant à obtenir une estimation des paramètres telle que la vraisemblance d'obtenir ses valeurs soit maximale ;
- Les moments, s'appuyant sur les moments de la distribution de l'échantillon de données (moyenne, variance, skewness⁵, kurtosis⁶).

Rappelons que l'étude menée ici cherche à évaluer les actifs en valeur de marché, sur une longue période pour tenir compte de la durée des engagements des assureurs. Or réaliser une évaluation en valeur de marché ne signifie pas uniquement prendre en compte les données de l'année ou proches de la date d'évaluation. En effet, pour la modélisation du taux court, nous utiliserons un historique de 30 ans minimum, ceci évite d'intégrer des erreurs si les deux dernières années avaient un comportement atypique par rapport à l'ensemble d'une courbe.

Comme on le verra dans la mise en œuvre, nous travaillerons principalement sur les données de Friggit qui sont annuelles, la série à étudier comporte donc entre 30 et 50 valeurs, ce qui forme un petit échantillon. Le calibrage par la méthode des moments n'est pas très adapté, en effet pour trouver la distribution de données en adéquation avec l'échantillon, il faudrait des séries de données importantes.

a) Méthode des moindres carrés ordinaires

La méthode des moindres carrés présentée ici est basée sur une régression linéaire. Nous avons à disposition un échantillon constitué de n couples d'observations (x_i, y_i) . A partir de ces éléments, on cherche la droite qui modélise le mieux le nuage de points. Cette droite doit être telle que la somme des carrés des écarts entre les y observés et les y théoriques soit minimale. L'estimation doit donc permettre de minimiser les résidus.

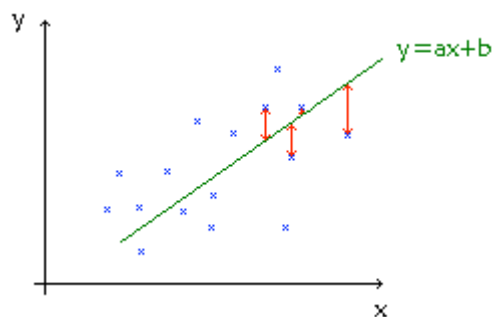


Fig : y observés – y théoriques – résidus à minimiser

La droite de régression cherchée est : $y_i = \alpha x_i + \beta + \varepsilon_i$, où y_i est la variable à expliquer, x_i est la variable explicative. L'enjeu de la régression est d'estimer les paramètres α et β (ici $p = 2$), tout en minimisant les résidus, noté ε_i .

L'équation peut aussi s'écrire sous forme matricielle, ce qui donne : $y = X\theta + \varepsilon$

⁵ Skewness : mesure l'asymétrie de la densité de probabilité d'une variable aléatoire ;

⁶ Kurtosis : mesure l'aplatissement de la densité de probabilité d'une variable aléatoire.

$$\text{Avec } y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \theta = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}.$$

L'estimation des paramètres par les MCO s'obtient par la formule : $\hat{\theta} = (X'X)^{-1}X'y$.

Les paramètres estimés $\hat{\alpha}$ et $\hat{\beta}$ sont tels que le SCR (Somme des Carrés des Résidus) soit minimal : $SCR = \sum_i (y_i - \hat{y}_i)^2$, y_i étant les valeurs observées et $\hat{y}_i$ les valeurs estimées.

Ou encore, la méthode minimise le CMR (Carré Moyen des Résidus), qui correspond au SCR corrigé des degrés de liberté (dit aussi sans biais), noté ddl : $CMR = \frac{\sum_i (y_i - \hat{y}_i)^2}{ddl}$ avec $ddl = n - p - 1$.

Par ailleurs, ce modèle est contraint par de fortes hypothèses qui doivent être respectées pour que l'estimation ait un sens :

- $\varepsilon \sim N(0, \sigma^2)$, soit :
 - ε a une distribution normale
 - $E(\varepsilon_i) = 0 \quad \forall i$
- ε_i sont *iid* (indépendamment et identiquement distribués)
 - L'indépendance correspond à la non autocorrélation des erreurs : $Var(\varepsilon_i, \varepsilon_j) = 0 \quad \forall i \neq j$
 - Identiquement Distribué, signifie que la propriété d'homoscédasticité doit être vérifiée : $Var(\varepsilon_i) = \sigma^2 \quad \forall i$

b) Méthode des moindres carrés non linéaires

Le modèle de régression non linéaire (*dans les paramètres*) peut être présenté de la manière suivante :

$$y_i = f(\theta, x_i) + \varepsilon_i$$

Avec y_i est la variable à expliquer,

x_i la variable explicative,

θ le vecteur des p paramètres à estimer.

Par ailleurs, ce modèle est contraint par les mêmes hypothèses que les MCO, à savoir :

- $\varepsilon \sim N(0, \sigma^2)$, soit :
 - ε a une distribution normale
 - $E(\varepsilon_i) = 0 \quad \forall i$
- ε_i sont *iid* (indépendamment et identiquement distribués)
 - L'indépendance correspond à la non autocorrélation des erreurs : $Var(\varepsilon_i, \varepsilon_j) = 0 \quad \forall i \neq j$
 - Identiquement Distribué, signifie que la propriété d'homoscédasticité doit être vérifiée : $Var(\varepsilon_i) = \sigma^2 \quad \forall i$

La méthode des moindres carrés non linéaires consiste à minimiser le carré des écarts entre les valeurs de y et les valeurs prises par la fonction f . Pour estimer le vecteur de paramètres θ , on minimise :

$$\text{Min}_{\theta} \text{SCE}(\theta) = \text{Min}_{\theta} \sum_{i=1}^T (y_i - f(\theta, x_i))^2 = r(\theta)'r(\theta)$$

avec r la fonction des résidus.

Pour estimer les paramètres plusieurs méthodes sont possibles :

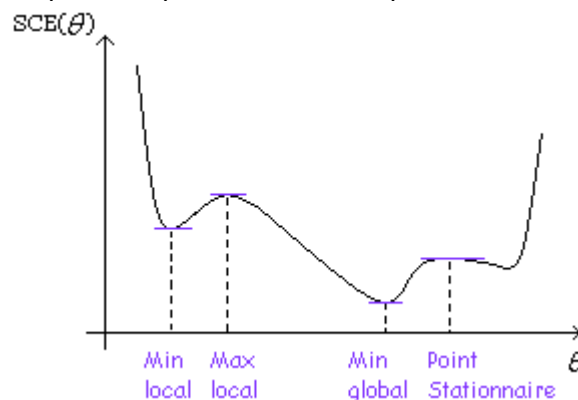
Le gradient

Pour trouver le minimum la condition de premier ordre doit être remplie, à savoir le gradient⁷ doit s'annuler : $\frac{\partial \text{SCE}(\theta)}{\partial \theta_i} = 0$.

$$\frac{\partial \text{SCE}(\theta)}{\partial \theta_i} = -2 \sum_{i=1}^T (y_i - f(\theta, x_i)) \frac{\partial f(\theta, x_i)}{\partial \theta_i}$$

Cette condition n'est pas suffisante pour s'assurer d'être à un minimum global. En effet lorsque les conditions de premier ordre sont vérifiées, on peut aussi se trouver sur un minimum local, un maximum local ou un point stationnaire.

Illustration des possibilités pour lesquels les dérivées premières s'annulent :



Avec la matrice Hessienne, qui est composée des dérivées secondes, il est possible de déterminer la nature du point dit « critique », elle est définie de la façon suivante :

$$H(\theta) = \begin{bmatrix} \frac{\partial^2 \text{SCE}(\theta)}{\partial a^2} & \frac{\partial^2 \text{SCE}(\theta)}{\partial a \partial b} \\ \frac{\partial^2 \text{SCE}(\theta)}{\partial b \partial a} & \frac{\partial^2 \text{SCE}(\theta)}{\partial b^2} \end{bmatrix}$$

Si θ est un minimum local, la hessienne $H(\theta)$ sera définie positive⁸.

Lorsque les conditions de premier et second ordre sont vérifiées, on se situe soit sur un minimum local, soit un minimum global. Il faut ensuite partir de différents points de départ : si on arrive toujours au même minimum, il est probable que ce soit le minimum global. Cette méthode permet de trouver facilement un minimum local, mais qui n'est pas nécessairement global.

⁷ Gradient : vecteur ayant pour composante les dérivées partielles d'une fonction.

⁸ Matrice définie positive : une matrice symétrique A est définie positive si pour tout vecteur x non nul, on a : $x'Ax > 0$

Un autre inconvénient de la méthode du gradient, c'est qu'elle ne permet pas toujours de trouver les valeurs qui minimisent la fonction $SCE(\theta)$. Comme f est une fonction non-linéaire des paramètres, il n'existe pas de solution fermée à l'équation (c'est-à-dire que la solution ne s'exprime pas avec $+$ / $-$ * ni exp ou log, mais plus certainement comme une limite, une intégrale, etc...)...c'est pourquoi il est préférable d'utiliser un algorithme qui soit plus performant et robuste que la méthode du gradient.

Pour faciliter l'obtention de θ , nous utilisons un algorithme de minimisation.

Algorithme par itérations successives

Une alternative à la méthode du gradient est un algorithme qui procède par itérations successives. Pour réaliser cette étape il existe différents algorithmes, dont les plus connus sont ceux de :

- Gauss - Newton
- Levenberg – Marquardt
- Newton

Le principe général de ce type d'algorithme est de choisir un point de départ, de déterminer la direction de pente de la descente de la fonction de coût $SCE(\theta)$, puis d'effectuer un pas en ce sens. Il faut répéter l'opération jusqu'à obtention du minimum, respectant le critère d'arrêt défini au préalable.

Les méthodes se différencient par le procédé de détermination de la pente et le choix du pas (constant, variable, etc...).

Le détail de l'algorithme de Gauss-Newton est donné en annexe (cf. §Algorithme de GAUSS - NEWTON).

c) Méthode du maximum de vraisemblance

L'objectif de cette méthode est de trouver les paramètres qui permettent de maximiser la vraisemblance d'obtenir les valeurs de l'échantillon de données.

Cette méthode s'appuie sur la définition de densité de probabilité de la variable aléatoire.

Nous disposons d'un échantillon de n données indépendantes : $X_1, X_2, \dots, X_n$.

X est une variable aléatoire de densité de probabilité $f(x, \theta)$ où θ est un vecteur de p paramètres, alors la fonction du max de vraisemblance est donnée par :

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

Elle représente la densité jointe de x_i (comme les x_i sont indépendants, il suffit de faire le produit des densités), c'est-à-dire qu'on obtient des renseignements sur la probabilité d'observer l'échantillon des n x_i .

L'objectif est de trouver la valeur de θ qui maximise la fonction de vraisemblance $L(\theta)$.

En pratique, on maximise le logarithme de la fonction de vraisemblance :

$$\log(L(\theta)) = \sum_{i=1}^n f(x_i; \theta)$$

Il est équivalent de maximiser la log vraisemblance ou la vraisemblance car la fonction logarithme est croissante. De plus, il est plus facile de manipuler des sommes que des produits pour les dérivées.

Un maximum de la fonction de vraisemblance est atteint lorsque :

- Les dérivées premières s'annulent
- Les dérivées secondes sont négatives ou nulles (matrice hessienne définie négative au point θ)

2. Matrice de variance – covariance / corrélation

Pour avoir une idée plus précise de la qualité de l'estimateur, il est nécessaire de déterminer la matrice de variance – covariance. Elle permet d'identifier les différentes interactions entre les paramètres estimés.

Si les paramètres sont indépendants, seule la diagonale de la matrice de variance-covariance des résidus sera non nulle.

Une covariance faible signifie que les variables explicatives sont faiblement liées, leurs évolutions sont relativement indépendantes. En revanche, une covariance forte sera le signe d'une liaison importante (dépendance) entre les variables explicatives.

Un coefficient positif indiquera que les variables évoluent dans la même direction. A contrario, un coefficient négatif signifie que les variables évoluent de façon opposée.

$$\Sigma_X = \text{var} \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} \text{var}(X_1) & \text{cov}(X_1 X_2) & \cdots & \text{cov}(X_1 X_p) \\ \text{cov}(X_2 X_1) & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \text{cov}(X_p X_1) & \cdots & \cdots & \text{var}(X_p) \end{pmatrix} = \begin{pmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} & \cdots & \sigma_{x_1 x_p} \\ \sigma_{x_2 x_1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \sigma_{x_p x_1} & \cdots & \cdots & \sigma_{x_p}^2 \end{pmatrix}$$

Avec

$$\begin{aligned} \sigma_{x_i x_j} &= E[(X_i - E(X_i))(X_j - E(X_j))] \\ &= E(X_i X_j) - E(X_i)E(X_j) \end{aligned}$$

Cette matrice correspond simplement à la matrice de variance covariance réduite des écarts-type. Elle indique l'influence des variables les unes par rapport aux autres. A la différence de la matrice de covariance, les coefficients ne sont pas sensibles aux changements d'unités des variables, il est donc plus facile d'interpréter et d'étudier les coefficients de la matrice de corrélation.

$$\text{Corr}(X_i, X_j) = \frac{\text{Cov}(X_i, X_j)}{\sqrt{\text{Var}(X_i)}\sqrt{\text{Var}(X_j)}}$$

Ce qui donne :

$$\text{Corr}(X_i, X_j) = \begin{pmatrix} 1 & \text{corr}(X_1 X_2) & \cdots & \text{corr}(X_1 X_p) \\ \text{corr}(X_2 X_1) & 1 & & \vdots \\ \vdots & & 1 & \vdots \\ \text{corr}(X_p X_1) & \cdots & \cdots & 1 \end{pmatrix}$$

En général, il est préférable que les variables ne soient pas corrélées, donc que les corrélations soient proches de zéro. Pour autant, avoir des coefficients proches de un ne signifie pas que le modèle est faux, particulièrement dans le cas de processus qui sont des modèles particuliers.

Remarque : la diagonale est uniquement constituée de « 1 » car les variables sont toutes parfaitement corrélées avec elles-mêmes.

a) Méthode des moindres carrés non linéaires

Dans le cas de la méthode d'estimation par les moindres carrés non linéaires, il n'est pas possible de calculer de façon exacte les valeurs de cette matrice, en revanche il existe des méthodes d'estimation.

La formule suivante donne une estimation de cette matrice pour la méthode des moindres carrés non linéaires :

$$Cov(\theta) = \frac{r(\theta)'r(\theta)}{n-p} * H(\theta)^{-1}$$

Où p le nombre de paramètres à estimer et $mse = \frac{r(\theta)'r(\theta)}{n-p}$.

Pour plus de détail sur cette propriété, vous pouvez vous référer à l'article de Gallant de 1975.

b) Méthode du maximum de vraisemblance

Dans le cas de la méthode d'estimation par le maximum de vraisemblance, il n'est pas non plus possible de calculer de façon exacte les valeurs de cette matrice, en revanche il existe des méthodes d'estimation.

Une alternative consiste à utiliser une méthode d'estimation basée sur la matrice Hessienne.

La formule suivante donne une estimation de cette matrice :

$$Cov(\theta) = -H(\theta)^{-1}$$

Preuve :

On démontre l'égalité : $Cov(\theta) = I(\theta)^{-1}$, où $I(\theta)$ est la matrice d'information de Fisher.

La matrice d'information de Fisher est définie comme la variance associée à la valeur θ du maximum de vraisemblance :

$$I(\theta) = -\frac{\partial^2}{\partial \theta \partial \theta'} l(\theta) = -\sum_{i=1}^n \left(\frac{\partial^2}{\partial \theta \partial \theta'} \log(f(x_i; \theta)) \right) = -H(\theta)$$

Pour une démonstration complète, référez vous au chapitre 8 du livre « Estimation et inférence en économétrie » de Russell Davidson et James G. MacKinnon.

3. Ecart-type des estimateurs

L'écart type des estimateurs se déduit simplement en prenant la racine carrée des éléments de la diagonale de la matrice de variance-covariance.

Cet indicateur permet de connaître la dispersion de la variable autour de sa moyenne.

B. Qualité de l'ajustement du processus

Lorsque l'on réalise l'estimation des paramètres d'un modèle, il est essentiel de connaître la qualité de l'ajustement obtenu. Pour cela il existe plusieurs méthodes pouvant nous donner des indications :

- L'analyse de la variance (ANOVA) pour la significativité du modèle global
- Le coefficient de détermination ajusté ($R^2_{ajusté}$) pour l'adéquation du modèle aux données
- Le T-Student pour la significativité de chacun des estimateurs

1. Analyse de la variance (ANOVA)

L'analyse de la variance des valeurs estimées permet d'avoir une vision de la qualité statistique du calibrage (cohérence globale du modèle).

La variance cherche à mesurer les écarts par rapport à la moyenne de l'échantillon. En effet, pour un modèle de la forme $y_i = \alpha x_i + \beta + \varepsilon_i$, si toutes les valeurs de y étaient les mêmes, elles seraient égales à leur moyenne, c'est la raison pour laquelle les variations de y sont mesurées par rapport à sa moyenne $\bar{y}$.

Une mesure convenable de la variation des y est donnée par la somme des carrés totale (SCT) :

$$SCT(\theta) = \sum_{i=1}^n (y_i - \bar{y})^2 \text{ avec } \theta = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

Le but est ensuite de décomposer la SCT en deux parties, la première due à la variabilité du modèle, notée SCE, et la seconde à la variabilité des termes aléatoires ε , notée SCR.

On a alors $SCT(\theta) = SCE(\theta) + SCR(\theta)$

$$\text{avec } SCE(\theta) = \sum_{i=1}^n (\hat{y} - \bar{y})^2$$

$$\text{et } SCR(\theta) = \sum_{i=1}^n (y - \hat{y})^2 = \sum_{i=1}^n (\varepsilon'_i)^2$$

Preuve :

Il est possible d'utiliser ici le théorème de Pythagore car les résidus sont orthogonaux aux paramètres estimés.

$$\sum_{i=1}^n \|y - \bar{y}\|^2 = \sum_{i=1}^n \|y - \hat{y}\|^2 + \sum_{i=1}^n \|\hat{y} - \bar{y}\|^2$$

L'analyse de la variance est souvent synthétisée sous forme d'une table d'ANOVA, terme signifiant **AN**alysis **O**f **V**ariance, elle se présente de la façon suivante :

Source de variation	Variation	Degré de liberté (ddl)	Erreur quadratique (carré moyen)	Fisher
Modèle	SCE	p-1	$SCE / (p - 1)$	$F = \frac{SCE / (p - 1)}{SCR / (n - p)}$
Résidus	SCR	n-p	$SCR / (n - p)$	
Total	SCT	n-1	$SCT / (n - 1)$	

Notations : n = nombre d'observations,
p = nombre de paramètres estimés.

Les estimations des variations (SCE, SCR, SCT) sont biaisées⁹, pour obtenir les valeurs non biaisées, le carré des écarts (aussi nommé erreur quadratique) est corrigé des degrés de liberté.

Le nombre de degrés de liberté, noté ddl, représente le nombre d'observations de l'échantillon auquel on retranche le nombre de valeurs qu'on fixe dans notre modèle. Par exemple pour SCT, $ddl=(n-1)$ car au préalable il a fallu calculer la moyenne des observations, il n'y a donc plus que $n-1$ données disponibles.

L'ANOVA s'applique sous certaines conditions, en particulier l'indépendance des formes quadratiques.

Interprétation ANOVA

Le modèle sera d'autant plus performant qu'une grande part de la variance totale sera expliquée par le modèle (SCE grand) et que la variance expliquée par les résidus sera faible (SCR petit).

2. Test Fisher Snedecor (inclus dans ANOVA)

Fisher et Snedecor ont mis au point un test d'hypothèse (cf. §Principe des tests d'hypothèse) qui permet d'étudier la loi liée au rapport de deux formes quadratiques¹⁰ indépendantes.

Dans le test du tableau d'ANOVA, le but est de comparer deux variances qui sont issues du même échantillon. Pour cela nous avons les éléments du SCE et du SCR.

1) Forme quadratique 1 : $\frac{SCE(\theta)}{p-1} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{p-1} \quad \forall i \text{ de } 1 \text{ à } n, \text{ dont la variance réelle est } \sigma_{SCE}$

2) Forme quadratique 2 : $\frac{SCR(\theta)}{n-p} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-p} = \frac{\sum_{i=1}^n (\varepsilon'_i)^2}{n-p} \quad \forall i \text{ de } 1 \text{ à } n, \text{ dont la variance réelle est } \sigma_{SCR}$

On va tester l'hypothèse H_0 contre l'hypothèse H_1 :

$$\begin{cases} H_0 : \sigma_{SCE} = \sigma_{SCR} \\ H_1 : \sigma_{SCE} \neq \sigma_{SCR} \end{cases}$$

La statistique de test est le ratio F. On montre que F suit une loi de Fisher à $(p-1, n-p)$ degrés de liberté :

$$F = \frac{S_{SCE}^2}{S_{SCR}^2} = \frac{SCE/(p-1)}{SCR/(n-p)}$$

⁹ Biais : un estimateur est dit non biaisé si l'espérance du paramètre estimé est égale à sa vraie valeur, pour toutes les valeurs de l'échantillon : $E(\hat{\theta}) = \theta, \forall n$.

¹⁰ Forme quadratique : polynôme homogène de degré deux avec un nombre quelconque de variables (exemple : $f(x,y,z)=x^2+yz+z^2$).

Pour conclure le test, il est nécessaire de comparer le résultat obtenu pour $F_{\text{calculé}}$ avec la valeur seuil, F_{seuil} , issue de la table de Fisher pour le risque α choisi :

- Si $F_{\text{calculé}} < F_{\text{seuil}}$ l'hypothèse de H_0 est rejetée au risque d'erreur α , ce qui signifie que les variances sont différentes et que le modèle n'est pas significatif.
- Si $F_{\text{calculé}} \geq F_{\text{seuil}}$ l'hypothèse de H_0 ne peut pas être rejetée au risque d'erreur α . Le modèle n'est pas rejeté, ce qui ne veut pas dire qu'il est adéquat, mais il est suffisamment significatif pour poursuivre l'étude.

3. Coefficients de déterminations (R^2 et $R^2_{\text{ajusté}}$)

Le coefficient de détermination, noté R^2 , est un indicateur qui permet de vérifier l'adéquation entre le modèle et les données observées. Il est défini par le ratio entre la variance expliquée par le modèle et la variance totale, soit :

$$R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT}$$

La valeur de R^2 est comprise entre 0 et 1, plus le R^2 est proche de 1 et meilleur sera le modèle. Lorsqu'il est supérieur à 0.75, le modèle est jugé correct.

Attention toutefois à l'interprétation du R^2 , en effet cet estimateur est biaisé, il est préférable d'utiliser le R^2 ajusté des degrés de liberté (c'est-à-dire de tenir compte du nombre de variables). La raison est toute simple, plus notre modèle contiendra de variables explicatives et plus le R^2 sera élevé « de façon mécanique », ce qui ne signifiera pas pour autant que le modèle sera meilleur.

Le R^2 ajusté est donné par la formule :

$$R^2_{\text{ajusté}} = 1 - \frac{SCR/(n-p)}{SCT/(n-1)} = 1 - \frac{n-1}{n-p}(1-R^2) = 1 - \frac{n-1}{n-p}\left(1 - \frac{SCM}{SCT}\right)$$

Remarque : le R^2 ajusté n'est pas un vrai « carré » de ce fait, il peut arriver qu'il soit négatif.

4. Test de Student

Le test T-Student sert en général à la comparaison de la moyenne observée et de la moyenne réelle. Ici il va nous permettre de tester la significativité de chacun des estimateurs du modèle.

On admet qu'un estimateur est significatif si sa valeur est différente de zéro, ce qui indique que l'estimateur explique une part du modèle.

Soit $\hat{\beta}$ l'estimateur à tester et $\hat{\sigma}_{\hat{\beta}}^2$ la variance sans biais (c'est-à-dire corrigée des degrés de liberté $n-1$) de l'estimateur. Cette variance estimée est souvent nommée « standard error ». On va tester l'hypothèse H_0 contre l'hypothèse H_1 :

$$\begin{cases} H_0 : \mu = 0 \\ H_1 : \mu \neq 0 \end{cases}$$

On détermine la statistique de test : $t = \frac{\hat{\beta} - \mu}{\frac{\hat{\sigma}_{\hat{\beta}}}{\sqrt{n}}}$ qui suit une loi de Student à $(n-p)$ degré de

liberté. La loi de Student est proche de la loi normale dans sa distribution, mais la loi de Student est plus appropriée dans le cas de petits échantillons (<30), ce qui est le cas dans l'étude présentée.

Avec l'hypothèse $H_0 : \mu = 0$, cela donne :

$$t_{\text{Calculé}} = \frac{\hat{\beta}}{\frac{\hat{\sigma}_{\hat{\beta}}}{\sqrt{n}}}$$

La conclusion du test se fait en vérifiant que la valeur $t_{\text{Calculé}}$ est dans l'intervalle

$\mu - t_{1-\alpha/2}(n-p) * \frac{\hat{\sigma}_{\hat{\beta}}}{\sqrt{n}} \leq t_{\text{Calculé}} \leq \mu + t_{1-\alpha/2}(n-p) * \frac{\hat{\sigma}_{\hat{\beta}}}{\sqrt{n}}$ avec t qui suit une loi de Student à $n-p$ degrés de liberté, pour un risque α donné :

- $t_{\text{Calculé}}$ n'est pas dans l'intervalle, alors l'hypothèse H_0 est rejetée au risque d'erreur α , $\hat{\beta}$ a une espérance différente de 0 ;
- $t_{\text{Calculé}}$ est dans l'intervalle, alors l'hypothèse H_0 est acceptée, $\hat{\beta}$ a une espérance nulle.

C. Test des hypothèses du modèle

A chaque modèle utilisé, un certain nombre d'hypothèses sont associées, il est essentiel de vérifier la justesse de ces hypothèses, notamment sur les résidus.

Lorsque les hypothèses ne sont pas vérifiées, les valeurs estimées ont de grandes chances d'être fausses .

Dans le modèle d'estimation par les moindres carrés non linéaires et du maximum de vraisemblance, les résidus doivent vérifier les caractéristiques suivantes :

- Suivre une loi normale centrée réduite
- Être homoscédastiques, c'est-à-dire de variance constante
- Être *iid* (indépendamment et identiquement distribués), c'est-à dire non autocorrélés

Des tests d'hypothèses peuvent être réalisés pour vérifier chacun de ces points.

1. Test de normalité des résidus

Dans la littérature, il existe différents tests permettant de vérifier l'adéquation d'un échantillon de données avec la distribution de la loi normale. Une partie d'entre eux peut être qualifiée de « graphique » car l'adéquation des données s'apprécie de façon visuelle, c'est le cas pour :

- L'histogramme des fréquences
- La boîte à moustache (*non décrite dans le document*)
- La droite de Henry (QQ-plot) (*non décrite dans le document*)

L'inconvénient de certaines de ces méthodes est qu'elles nécessitent de nombreuses données, or certaines de nos échantillons ne comportent qu'une trentaine de données !

Les coefficients d'asymétrie (Skewness) et d'aplatissement (Kurtosis) d'une distribution sont de bons indicateurs, ils correspondent aux moments d'ordre 3 et 4. Il suffit alors de les comparer à ceux de la loi normale (Skewness = 0, Kurtosis = 3) pour évaluer l'adéquation des données.

Enfin, des méthodes plus rigoureuses s'appuient sur des tests d'hypothèses. Plusieurs sont à notre disposition, dont les plus connus sont :

- Shapiro – Wilk : adéquat pour de petits échantillons (< 50) ;
- Jarque – Bera : d'autant plus puissant que l'échantillon sera grand ;
- Kolmogorov – Smirnov : adéquat pour de grands échantillons ;
- Lilliefors : variante de Kolmogorov – Smirnov avec moyenne et variance inconnues ;
- Anderson – Darling : variante de Kolmogorov – Smirnov donnant de l'importance aux queues de distributions ;
- Etc...

Notre étude se basant sur un petit échantillon de données, les tests de Kolmogorov – Smirnov et leurs dérivés ne sont pas adaptés. Nous allons nous intéresser aux tests de Shapiro – Wilk et Jarque – Bera.

a) Shapiro – Wilk

Le test de Shapiro-Wilk est basé sur la statistique de test suivante :

$$W = \frac{\left(\sum_{i=1}^{E[n/2]} a_i (e_{n-i+1} - e_i) \right)^2}{\sum_{i=1}^n (e_i - \bar{e})^2}$$

Avec n : nombre d'observations de l'échantillon ;

e_i : résidus du modèle ;

a_i : coefficients générés à partir de la moyenne et de la matrice de variance covariance des quantiles d'un échantillon de taille n suivant une loi normale (coefficients provenant d'une table spécifique au test de Shapiro Wilk) ;

$E[n/2]$: partie entière de $n/2$.

On pose l'hypothèse :

$$\begin{cases} H_0 : \text{La série est gaussienne} \\ H_1 : \text{La série n'est pas gaussienne} \end{cases}$$

La réalisation du test s'effectue en plusieurs étapes :

1. Tri des données de résidus par ordre croissant :

$$e_1 \leq e_2 \leq \dots \leq e_n$$

2. Calcul de la somme des carrés des écarts :

$$\sum_{i=1}^n (e_i - \bar{e})^2$$

3. Calcul du numérateur à l'aide des coefficients a_i de la table spécifique Shapiro-Wilk :

$$\left(\sum_{i=1}^{E[n/2]} a_i (e_{n-i+1} - e_i) \right)^2$$

4. Détermination de la statistique :

$$W_{\text{Calculé}} = \frac{\left(\sum_{i=1}^{E[n/2]} a_i (e_{n-i+1} - e_i) \right)^2}{\sum_{i=1}^n (e_i - \bar{e})^2}$$

5. Conclusion du test en comparant la valeur $W_{\text{Calculé}}$ à la valeur W_{seuil} pour un risque α donné :

- $W_{\text{Calculé}} > W_{\text{seuil}}$ alors l'hypothèse de normalité ne peut pas être rejetée, la normalité de l'échantillon est acceptée ;
- $W_{\text{Calculé}} \leq W_{\text{seuil}}$ alors l'hypothèse de normalité est rejetée.

b) Jarque – Bera

Le test de Jarque - Bera se base sur les coefficients de Skewness et de Kurtosis.

La Skewness mesure le degré d'asymétrie d'une distribution, elle correspond au moment d'ordre 3 de la distribution suivi par la variable aléatoire X , soit :

$$\text{CoefS} = \frac{E(X - \mu)^3}{\left(\sqrt{E(X - \mu)^2} \right)^3}$$

Un coefficient de skewness égal à 0 signifie que la distribution est symétrique. S'il est inférieur à 0, l'asymétrie sera vers la gauche et, inversement, vers la droite s'il est positif.

La Kurtosis mesure le degré d'aplatissement d'une distribution, elle correspond au moment d'ordre 4 de la distribution suivi par la variable aléatoire X , soit :

$$\text{CoefK} = \frac{E(X - \mu)^4}{\left(\sqrt{E(X - \mu)^2} \right)^2}$$

Un coefficient de kurtosis à 3 correspond à celui d'une loi normale. Si le coefficient se rapproche de 0, c'est que la distribution est plus aplatie, et inversement.

L'indicateur Jarque-Bera est basé sur la statistique de test suivante :

$$JB = \frac{n-p}{6} \left(\text{CoefS}^2 + \frac{(\text{CoefK} - 3)^2}{4} \right)$$

L'hypothèse posée est la même que pour Shapiro-Wilk :

$$\begin{cases} H_0 : \text{La série est gaussienne} \\ H_1 : \text{La série n'est pas gaussienne} \end{cases}$$

La réalisation du test s'effectue en plusieurs étapes :

1. Centrer et réduire les résidus
2. Calculer les coefficients de Skewness et Kurtosis
3. Déterminer la statistique de test
4. Conclure sur le test en comparant la valeur $JB_{Calculé}$ à la valeur JB_{seuil} qui suit une loi de χ^2 à deux degrés de liberté, pour un risque α donné :
 - $JB_{Calculé} > JB_{seuil}$ alors l'hypothèse de normalité est rejetée, la normalité de l'échantillon est rejetée ;
 - $JB_{Calculé} \leq JB_{seuil}$ alors l'hypothèse de normalité ne peut pas être rejetée.

2. Test de normalité standardisé des résidus

a) Test de la moyenne égale à zéro

De la même façon que pour le test sur la significativité des paramètres, nous pouvons utiliser le test T-Student qui est un test de comparaison de moyenne.

La statistique de test t suit une loi de Student à (n-1) degrés de liberté :

$$t = \frac{m_r - \mu}{\frac{S_r}{\sqrt{n}}}$$

Avec m_r la moyenne de l'échantillon testé, donc des résidus ;

s_r^2 la variance de l'échantillon testé, donc des résidus.

On va tester l'hypothèse H_0 contre l'hypothèse H_1 :

$$\begin{cases} H_0 : \mu = 0 \\ H_1 : \mu \neq 0 \end{cases}$$

La réalisation du test s'effectue en plusieurs étapes :

1. Calculer la moyenne et la variance des résidus
2. Etablir la statistique de test
3. Conclure sur le test en vérifiant que la valeur $t_{Calculé}$ est dans l'intervalle

$$\mu - t_{1-\alpha/2}(n-1) * \frac{S_r}{\sqrt{n}} \leq t_{Calculé} \leq \mu + t_{1-\alpha/2}(n-1) * \frac{S_r}{\sqrt{n}}$$

à la valeur t qui suit une loi de Student à n-1 degrés de liberté, pour un risque α donné :

- $t_{Calculé}$ n'est pas dans l'intervalle, alors l'hypothèse H_0 est rejetée, la moyenne n'est pas μ ;
- $t_{Calculé}$ est dans l'intervalle, alors l'hypothèse H_0 ne peut pas être rejetée, la moyenne de l'échantillon vaut μ .

b) Test de la variance égale à un

Le test de comparaison d'une variance à une valeur peut s'appliquer si :

- L'échantillon testé est constitué de n observations qui sont indépendantes
- La distribution de données de l'échantillon suit une loi normale

La statistique de test V suit une loi de χ^2 à $(n-1)$ degrés de liberté :

$$V = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sigma_0^2}$$

Avec x_i les observations de l'échantillon

$\bar{x}$ la moyenne de l'échantillon

σ_0^2 la valeur donnée à la variance

On va tester l'hypothèse H_0 contre l'hypothèse H_1 :

$$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \neq \sigma_0^2 \end{cases}$$

La réalisation du test s'effectue en plusieurs étapes :

1. Calculer la moyenne et la variance des résidus
2. Calculer la somme des carrés des écarts des résidus
3. Établir la statistique de test
4. Conclusion du test en vérifiant que la valeur $V_{\text{Calculé}}$ est, pour un risque α donné, dans l'intervalle :

$$\chi_{\alpha/2}^2(n-1) \leq V_{\text{Calculé}} \leq \chi_{1-\alpha/2}^2(n-1)$$

- $V_{\text{Calculé}}$ est dans l'intervalle, alors l'hypothèse ne peut pas être rejetée, la variance est égale à σ_0^2 ;
- $V_{\text{Calculé}}$ n'est pas dans l'intervalle, alors l'hypothèse H_0 est rejetée.

3. Test d'indépendance des résidus (autocorrélation)

Dire qu'il existe une autocorrélation des résidus signifie que la covariance¹¹ des termes d'erreurs n'est pas nulle, donc que les termes d'erreurs sont dépendants au cours du temps. L'autocorrélation est principalement liée aux séries temporelles en raison de l'ordre chronologique des données, le terme d'erreur d'une date dépend du terme d'erreur à une autre date.

Une autocorrélation positive signifie que les erreurs successives sont de même signe (+++, ---, +++).

Lorsque l'on cherche à tester l'autocorrélation d'ordre 1, cela signifie que l'on cherche la corrélation d'une série avec elle-même décalée d'une période (ordre k = décalée de k périodes).

Différents tests existent dans la littérature :

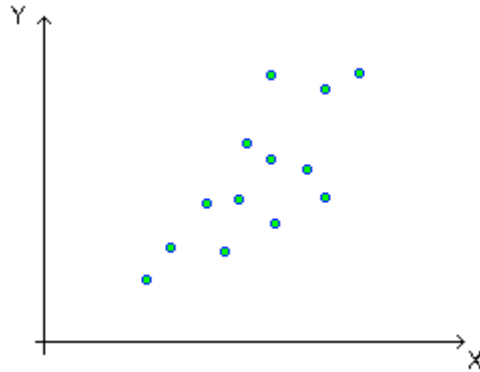
- Contrôle graphique,
- Test de Durbin & Watson : permet de tester l'autocorrélation uniquement à l'ordre 1, lorsque l'équation de régression ne contient pas d'endogène retardée ;
- Test de Breusch Godfrey : permet de tester l'autocorrélation à l'ordre p , même lorsque l'équation de régression contient une endogène retardée.

¹¹ Covariance : permet d'évaluer le sens de variation de deux variables pour qualifier leurs indépendances. Si $\text{cov}(x,y)=0$ alors les variables x et y sont indépendantes.

Dans cette étude il est préférable d'utiliser le test de Breusch Godfrey car l'équation de régression contient une endogène retardée comme variable explicative.

a) Contrôle graphique

L'utilisation d'un graphique permet de visualiser rapidement les autocorrélations lorsqu'elles sont flagrantes. Ci-dessous, un exemple d'autocorrélation :



Lorsque les résidus sont répartis de façon homogène dans le graphique, on peut considérer qu'il n'y a pas d'autocorrélation des résidus. Toutefois, il sera préférable de la vérifier avec des méthodes plus précises.

b) Breusch-Godfrey

Pour utiliser le test de Breusch-Godfrey (aussi appelé LM test), on fait l'hypothèse que les résidus sont de type autorégressif d'ordre p (AR (p)), c'est-à-dire qu'ils s'expriment sous la forme :

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \dots + \rho_p \varepsilon_{t-p} + u_t$$

Avec $u_t \sim \text{i.i.d.}$;

$$E(u_t) = 0 ;$$

$$V(u_t) = \sigma^2.$$

Le test de Breusch Godfrey consiste à estimer les coefficients d'autocorrélation ρ par les moindres carrés à partir d'une régression auxiliaire sur les variables explicatives du modèle et les résidus retardés.

On va tester l'hypothèse H_0 contre l'hypothèse H_1 :

$$\begin{cases} H_0 : \rho_1 = \rho_2 = \dots = \rho_p = 0 \\ H_1 : (\rho_1, \rho_2, \dots, \rho_p) \neq (0, 0, \dots, 0) \end{cases}$$

Si H_0 est vérifiée, il n'y a pas d'autocorrélation.

La réalisation du test s'effectue en plusieurs étapes :

1. Réaliser une estimation par les moindres carrés du modèle de base pour obtenir les résidus estimés $\hat{\varepsilon}_t$.
2. Régresser les résidus estimés sur les variables explicatives d'origine et les résidus décalés ($\varepsilon_{t-1}, \varepsilon_{t-2}, \text{etc...}$) à l'aide du modèle auxiliaire :

$$\varepsilon_t = \alpha_0 + \sum_{i=1}^m (\alpha_i x_i) + \sum_{k=1}^p (\rho_k \varepsilon_{t-k}) + z_t, \quad z_t \sim i.i.d(0, \omega^2)$$

Avec m le nombre de variables explicatives du modèle ;
 p l'ordre du modèle autorégressif.

3. Construire la statistique de test du multiplicateur de Lagrange (en supposant que les ε_t sont normalement distribués) : $LM = nR^2$ avec

$$R^2 = 1 - \frac{\sum_{t=p+1}^n (\hat{z}_t)}{\sum_{t=p+1}^n (\varepsilon_t)}$$

Remarque : Breusch a montré qu'asymptotiquement il y avait n degrés de liberté. En effet, une confusion intervient souvent, considérant qu'il y a $n-p$ degrés de liberté en raison de l'ordre p de retard. En réalité, on ne perd pas un degré car la valeur supprimée est remplacée par des zéros.

4. Conclure le test en comparant la valeur $LM_{Calculé}$ à la valeur LM_{seuil} qui suit une loi de Chi2 à p degrés de liberté ($LM_{seuil} \sim \chi^2(p)$), pour un risque α donné :
 - $LM_{Calculé} > LM_{seuil}$ alors l'hypothèse H_0 est rejetée, il existe une autocorrélation des résidus ;
 - $LM_{Calculé} \leq LM_{seuil}$ alors l'hypothèse H_0 est acceptée, les résidus sont indépendants.

4. Test sur l'hypothèse des résidus identiquement distribués

Engle a introduit en 1982 les modèles ARCH (AutoRegressive Conditionally Heteroscedastic) pour tenir compte du fait que l'erreur de prévision peut dépendre pour certaines séries de la taille de l'erreur de prévision précédente.

Une présentation simple du modèle peut être donnée par l'équation :

$$y_t = \beta x_t + u_t$$

Avec $u_t = v_t \sqrt{\alpha_0 + \alpha_1 u_{t-1}^2}$;
 $v_t \sim N(0,1)$;
 $\alpha_0 > 0, \alpha_1 \geq 0$;

On s'intéresse ensuite à l'espérance conditionnelle et à la variance conditionnelle pour identifier si le processus u_t impacte l'espérance et la variance du modèle.

Raisonnons sur un modèle ARCH(1), c'est-à-dire avec un retard d'ordre 1 :

$E(u_t / u_{t-1}) = E_{t-1}(v_t) * E_{t-1}(\alpha_0 + \alpha_1 u_{t-1}^2)^{1/2} = 0$ ce qui s'explique par le fait que u_{t-1}, v_t sont indépendants et $E(v_t) = 0$.

$$Var(u_t / u_{t-1}) = E[(u_t - E(u_t / u_{t-1}))^2 / u_{t-1}] \text{ or } E(u_t / u_{t-1}) = 0$$

$$\begin{aligned} Var(u_t / u_{t-1}) &= E[u_t^2 / u_{t-1}] \\ &= E_{t-1}[v_t^2] * E_{t-1}[\alpha_0 + \alpha_1 u_{t-1}^2] \\ &= \alpha_0 + \alpha_1 E_{t-1}[u_{t-1}^2] \\ &= \alpha_0 + \alpha_1 u_{t-1}^2 \end{aligned}$$

En généralisant à un modèle ARCH(q), le moment de second ordre est donné par l'expression :

$$E[u_t^2 / u_{t-1}^2, u_{t-2}^2, \dots, u_{t-q}^2] = \alpha_0 + \alpha_1 u_{t-1}^2 + \dots + \alpha_q u_{t-q}^2$$

Afin de vérifier que le processus suivi par les résidus n'ait pas d'influence sur les moments, il est nécessaire que les paramètres α_1 à α_q soient nuls. Ce qui impliquera que les résidus soient identiquement distribués. La variance conditionnelle sera alors constante et égale à α_0 .

On va tester l'hypothèse H_0 contre l'hypothèse H_1 :

$$\begin{cases} H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_q = 0 \\ H_1 : (\alpha_1, \alpha_2, \dots, \alpha_q) \neq (0, 0, \dots, 0) \end{cases}$$

Si H_0 est vérifiée, il y a identique distribution des résidus.

La réalisation du test s'effectue en plusieurs étapes :

1. Réaliser une estimation par les moindres carrés du modèle de base pour obtenir les résidus estimés $\hat{u}_t$.
2. Régresser les résidus au carré estimés sur la constante et les résidus au carré estimés retardés ($u_{t-1}^2, u_{t-2}^2, \dots$) à l'aide du modèle auxiliaire :

$$u_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \dots + \alpha_q u_{t-q}^2 + z_t, \quad z_t \sim i.i.d(0, \omega^2).$$

Avec q l'ordre du modèle autorégressif.

3. Construire la statistique de test du multiplicateur de Lagrange (en supposant que les u_t sont normalement distribués) : $LM = nR^2$ avec

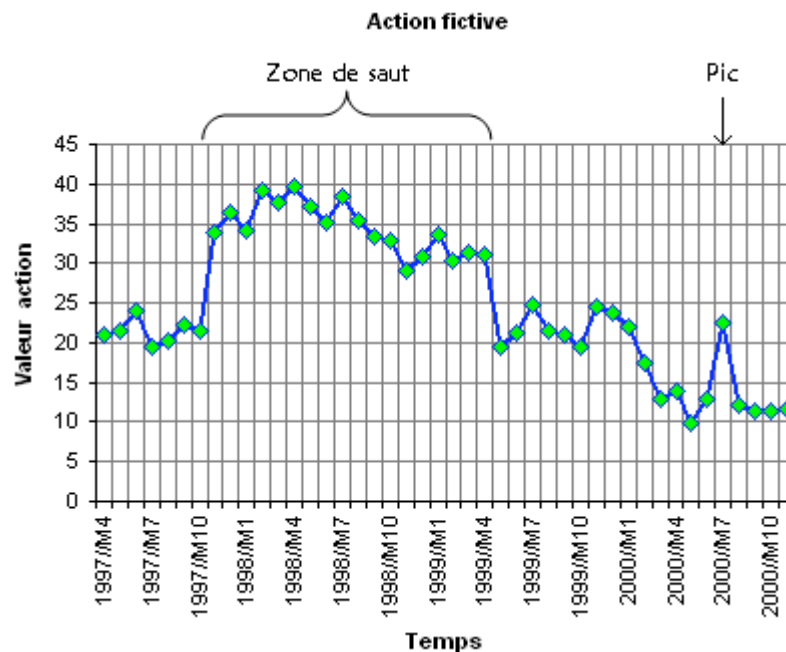
$$R^2 = 1 - \frac{\sum_{t=q+1}^n (\hat{z}_t)}{\sum_{t=q+1}^n (u_t)}$$

4. Conclure le test en comparant la valeur $LM_{Calculé}$ à la valeur LM_{seuil} qui suit une loi de χ^2 à q degrés de liberté ($LM_{seuil} \sim \chi^2(q)$), pour un risque α donné :
 - $LM_{Calculé} > LM_{seuil}$ alors l'hypothèse H_0 est rejetée, les résidus ne sont pas identiquement distribués ;
 - $LM_{Calculé} \leq LM_{seuil}$ alors l'hypothèse H_0 est acceptée, les résidus sont identiquement distribués.

D. Particularité d'un modèle intégrant des sauts : Merton

La modélisation des taux selon le modèle de Vasicek (un ou deux facteurs) ne traduit pas nécessairement correctement l'enjeu des actions. Ces produits sont risqués et possèdent donc des queues de distribution épaisses, ce qui peut être traduit dans les modèles par des « sauts ».

Attention, il ne faut pas confondre un « saut » dans la courbe avec un « pic » de valeur. Un saut est un déplacement du niveau moyen de la courbe pour une petite période, alors que pour un pic, il peut y avoir un point unique.



1. Présentation du modèle

En 1976, Merton complète le modèle de Black et Scholes en introduisant une composante à saut, avec l'utilisation d'un processus de Poisson.

Il est considéré dans ce modèle qu'entre deux instants de saut, les trajectoires suivent la dynamique du modèle de Black & Scholes.

Dans ce cas la distribution des sauts sera gaussienne et centrée, ceci signifie que les sauts sont symétriques.

La dynamique

La dynamique suivie par le cours de l'action est :

$$S(t) = S_0 \exp \left[\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W(t) + \underbrace{\sum_{k=1}^{N(t)} (U(k))}_{\text{Saut_Merton}} \right]$$

Avec N un processus de Poisson d'intensité λ

U une suite de variables aléatoires indépendantes et identiquement distribuées selon la loi normale $N \sim (0, \Omega)$

Les processus sont N , U et W indépendants

On considère un pas de temps $\delta = t_i - t_{i-1}$ constant sur l'intervalle $[0, T]$.

Le rendement de l'action sur l'intervalle $\delta = t_i - t_{i-1}$ sera :

$$\frac{S(t_i)}{S(t_{i-1})} = \frac{\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)(t_i) + \sigma W(t_i) + \sum_{k=1}^{N(t_i)} (U(k))\right)}{\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)(t_{i-1}) + \sigma W(t_{i-1}) + \sum_{k=1}^{N(t_{i-1})} (U(k))\right)}$$

$$\frac{S(t_i)}{S(t_{i-1})} = \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)\delta + \sigma(W(t_i) - W(t_{i-1})) + \sum_{k=N(t_{i-1})}^{N(t_i)} (U(k))\right)$$

Le rendement logarithmique, est noté :

$$r_{t_i} = \ln\left(\frac{S(t_i)}{S(t_{i-1})}\right)$$

$$r_{t_i} = \left(\left(\mu - \frac{\sigma^2}{2}\right)\delta + \sigma(W(\delta)) + \sum_{k=N(t_{i-1})}^{N(t_i)} (U(k))\right)$$

Remarque : les rendements ainsi définis sont indépendamment et identiquement distribués, les rendements du pas de temps $\delta = t_i - t_{i-1}$ ne dépendent pas des pas de temps précédent.

Pour réaliser une estimation par le maximum de vraisemblance, il est nécessaire de déterminer la densité de probabilité $f_r(x)$ suivie par R :

$$f_r(x) = \frac{e^{-\lambda\delta}}{\sqrt{2\pi}} \left[\sum_{n=0}^{\infty} \left(\frac{(\delta\lambda)^n}{n! * \sqrt{(\delta\sigma^2 + n\sigma_u^2)}} * \exp\left(-\frac{\left(x - \left(\mu - \frac{\sigma^2}{2}\right)\delta\right)^2}{2 * (\delta\sigma^2 + n\sigma_u^2)}\right) \right) \right]$$

Remarque : Il est important de conserver le delta, principalement lorsque la périodicité sur laquelle le calibrage est basé est différente de la périodicité de projection.

Le détail des calculs menant à l'expression de la densité est donné en annexe (cf. §Densité du modèle de Merton).

Le calibrage du modèle de Merton va légèrement différer des modèles de Vasicek présentés auparavant. Il n'est pas judicieux d'utiliser directement la méthode du maximum de vraisemblance qui ne nous assurerait pas d'arriver à un maximum global, de plus le nombre d'itérations serait important.

Nous allons procéder en deux étapes, comme présenté dans le document « Un modèle de détection des sauts dans un processus de cours boursiers » de Boni, Genot, Momein, à savoir :

- 1) L'estimation des paramètres par la méthode des moments : ceci permettra d'avoir une bonne idée de la valeur finale à trouver ;
- 2) L'estimation par le maximum de vraisemblance : en initialisant les paramètres du modèle avec les valeurs de la méthode des moments, il suffit alors de trouver un maximum local, donc beaucoup moins coûteux en nombre d'itérations.

Dans le modèle de Merton, les quatre paramètres à estimer sont : $\mu, \sigma, \lambda, \sigma_u$.

2. Estimation par la méthode des moments

La méthode des moments s'appuie sur la loi des grands nombres, en cherchant à égaliser les moments théoriques avec les moments empiriques constatés sur l'échantillon à disposition, il est alors possible d'obtenir les valeurs des paramètres.

Si θ est le vecteur des paramètres cherchés, on peut avoir la formule théorique des moments à partir de la densité de probabilité suivie par les données étudiées $M(\hat{\theta})$, et on peut calculer les moments constatés m , et en déduire θ par $m = M(\hat{\theta})$.

On s'intéresse au moment centré d'ordre pair : 2, 4 et 6. Les moments d'ordre impair étant nuls car la distribution des variables aléatoires sont symétriques.

Dans le cas de Merton, on a : $E(r) = \left(\mu - \frac{\sigma^2}{2} \right) \delta$.

Le calcul des moments d'ordre pair donne :

$$m_{2k} = E[(r - m)^{2k}] = \frac{2k!}{2^k k!} \sum_{n=0}^{\infty} \frac{(\delta\lambda)^n e^{(-\lambda\delta)}}{n!} (\delta\sigma^2 + n\sigma_u^2)^k$$

Calculons les moments théoriques d'ordre 2, 4 et 6 :

Pour k=1 : $m_2 = Var(r) = \delta(\sigma^2 + \lambda\sigma_u^2)$

Pour k=2 : $m_4 = 3\delta^2\sigma^4 + 6\lambda\delta^2\sigma^2\sigma_u^2 + 3\lambda^2\delta^2\sigma_u^4 + 3\lambda\delta\sigma_u^4$

Pour k=3 : $m_6 = 15\delta^3\sigma^6 + 45\lambda\delta^3\sigma^4\sigma_u^2 + 45\lambda^2\delta^3\sigma^2\sigma_u^4 + 45\lambda\delta^2\sigma^2\sigma_u^4$
 $+ 15\lambda^3\delta^3\sigma_u^6 + 45\lambda^2\delta^2\sigma_u^6 + 15\lambda\delta\sigma_u^6$

Avec les trois moments, on a trois équations à trois inconnues qui nous permettent d'obtenir les valeurs estimées de $\{\sigma, \lambda, \sigma_u\}$, puis nous pouvons déduire μ .

Pour résoudre ce système qui n'est pas linéaire, il faut utiliser un algorithme du type Newton-Raphson qui permet de trouver le minimum d'une fonction. La minimisation revient à obtenir les valeurs qui annulent la dérivée première.

Les dérivées sont données par :

Cas k=1

$$\frac{\partial m_2}{\partial \sigma} = 2\delta\sigma ;$$

$$\frac{\partial m_2}{\partial \lambda} = \delta\sigma_u^2 ;$$

$$\frac{\partial m_2}{\partial \sigma_u} = 2\lambda\delta\sigma_u$$

Cas k=2

$$\frac{\partial m_4}{\partial \sigma} = 12\delta^2 \sigma^3 + 12\lambda\delta^2 \sigma \sigma_u^2 ;$$

$$\frac{\partial m_4}{\partial \lambda} = 6\delta^2 \sigma^2 \sigma_u^2 + 6\lambda\delta^2 \sigma_u^4 + 3\delta \sigma_u^4 ;$$

$$\frac{\partial m_4}{\partial \sigma_u} = 12\lambda\delta^2 \sigma^2 \sigma_u + 12\lambda^2 \delta^2 \sigma_u^3 + 12\lambda\delta \sigma_u^3$$

Cas k=3

$$\frac{\partial m_6}{\partial \sigma} = 90\delta^3 \sigma^5 + 90\lambda\delta^2 \sigma \sigma_u^4 + 180\lambda\delta^3 \sigma^3 \sigma_u^2 + 90\lambda^2 \delta^3 \sigma \sigma_u^4 ;$$

$$\frac{\partial m_6}{\partial \lambda} = 45\delta^3 \sigma^4 \sigma_u^2 + 90\lambda\delta^3 \sigma^2 \sigma_u^4 + 45\lambda^2 \delta^3 \sigma_u^6 + 45\delta^2 \sigma^2 \sigma_u^4 + 90\lambda\delta^2 \sigma_u^6 + 15\delta \sigma_u^2 ;$$

$$\frac{\partial m_6}{\partial \sigma_u} = 90\lambda\delta^3 \sigma^4 \sigma_u + 180\lambda^2 \delta^3 \sigma^2 \sigma_u^3 + 90\lambda^3 \delta^3 \sigma_u^5 + 180\delta^2 \lambda \sigma^2 \sigma_u^3 + 270\lambda \sigma_u^5$$

3. Estimation par le maximum de vraisemblance

L'objectif est de trouver les paramètres qui permettent de maximiser la vraisemblance d'obtenir les valeurs de l'échantillon de données.

Cette méthode s'appuie sur la définition de densité de probabilité de la variable aléatoire.

Nous disposons d'un échantillon de n données indépendantes : $X_1, X_2, \dots, X_n$.

X est une variable aléatoire de densité de probabilité $f(x, \theta)$ où θ est un vecteur de p paramètres, alors la fonction du maximum de vraisemblance est donnée par :

$$L(\theta) = \prod_{i=1}^n f(r_i; \theta)$$

$$L(\theta) = \prod_{i=1}^n f_r(r_i, \lambda, \mu, \sigma^2, \sigma_u^2) = \prod_{i=1}^n \frac{e^{-\lambda\delta}}{\sqrt{2\pi}} \left[\sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n! \sqrt{(\delta\sigma^2 + n\sigma_u^2)}} * \exp \left(-\frac{\left(r_i - \left(\mu - \frac{\sigma^2}{2} \right) \delta \right)^2}{2 * (\delta\sigma^2 + n\sigma_u^2)} \right) \right) \right]$$

Avec $\theta = [\lambda, \mu, \sigma^2, \sigma_u^2]$

En pratique, on maximise le logarithme de la fonction de vraisemblance :

$$\Lambda = \log(L(\theta)) = \sum_{i=1}^n \log f(r_i; \theta)$$

$$\log(L(\theta)) = \sum_{i=1}^n \frac{e^{-\lambda\delta}}{\sqrt{2\pi}} \left[\sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n! \sqrt{(\delta\sigma^2 + n\sigma_u^2)}} * \exp \left(-\frac{\left(r_i - \left(\mu - \frac{\sigma^2}{2} \right) \delta \right)^2}{2 * (\delta\sigma^2 + n\sigma_u^2)} \right) \right) \right]$$

$$\Lambda = N\left(-\lambda\delta - \log(\sqrt{2\pi})\right) + \sum_{i=1}^N \log \left[\sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n! \sqrt{(\delta\sigma^2 + n\sigma_u^2)}} * \exp \left(-\frac{\left(r_i - \left(\mu - \frac{\sigma^2}{2} \right) \delta \right)^2}{2 * (\delta\sigma^2 + n\sigma_u^2)} \right) \right) \right]$$

Remarque : lorsque $\sigma_u^2 = 0$, il n'y a plus de saut, on se ramène au modèle de Black et Scholes.

Pour maximiser la log vraisemblance, il faut calculer les dérivées premières. Les dérivées secondes vont être utiles pour déterminer la variance des estimateurs qui est : $Cov(\theta) = -H(\theta)^{-1}$

Pour faciliter l'écriture et la lecture du calcul, posons :

$$g(r_{t_i}) = \sum_{n=0}^{\infty} \left[\frac{(\lambda\delta)^n}{n! \sqrt{\delta\sigma^2 + n\sigma_u^2}} \exp \left(-\frac{\left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)^2}{2(\delta\sigma^2 + n\sigma_u^2)} \right) \right]$$

$$v(r_{t_i}) = -\frac{\left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)^2}{2(\delta\sigma^2 + n\sigma_u^2)}$$

$$\text{D'où } g(r_{t_i}) = \sum_{n=0}^{\infty} \left[\frac{(\lambda\delta)^n}{n! \sqrt{\delta\sigma^2 + n\sigma_u^2}} \exp(v(r_{t_i})) \right]$$

$$\text{Donc } \Lambda = -N\left(\lambda\delta + \log(\sqrt{2\pi})\right) + \sum_{i=1}^N \log(g(r_{t_i}))$$

Les calculs intermédiaires à réaliser sont dans l'ordre :

- Les dérivées premières de v ;
- Les dérivées secondes de v ;
- La déduction des dérivées premières de $g(r_{t_i})$;
- La déduction des dérivées secondes de $g(r_{t_i})$;
- Enfin, on obtient les dérivées premières de la log vraisemblance ;
- Et les dérivées secondes de la log vraisemblance.

Le détail des calculs est développé dans l'annexe (cf. § Estimation des paramètres par le maximum de vraisemblance).

4. Qualité d'ajustement du modèle et vérification des hypothèses

Le modèle de Merton est différent des autres modèles car il introduit deux variables aléatoires dans son processus, de ce fait il n'est pas possible d'utiliser les mêmes outils que pour les modèles de Vasicek pour vérifier la qualité et les hypothèses du modèle.

Comme l'on présenté Boni, Genot, Momein, il est possible d'utiliser :

- 1) Le test de la vraisemblance : consistant à évaluer le gain entre les vraisemblances du modèle intégrant le saut ou non ;
- 2) Le test d'adéquation du Chi2 : permettant de vérifier que les rendements ne suivent pas une loi normale, mais intègrent bien des sauts ;
- 3) Le test du up and down : vérifiant le caractère indépendant des rendements.

a) Test de la vraisemblance

Ce test est très simple, il s'agit de faire le rapport des vraisemblances entre le modèle de Merton n'intégrant pas le saut (= Black et Scholes) et celui prenant en compte les sauts, pour déterminer s'il est intéressant d'utiliser un modèle à saut.

On va tester l'hypothèse H_0 contre l'hypothèse H_1 :

$$\begin{cases} H_0 : (\lambda, \sigma_u) = (0, 0) \\ H_1 : (\lambda, \sigma_u) \neq (0, 0) \end{cases}$$

Si H_0 est vérifiée, le modèle à saut n'est pas adapté.

La statistique de test est donnée par :

$$L = \frac{\prod_{i=1}^n \frac{1}{\sigma \sqrt{2\pi}} \exp \left(-\frac{\left(r_i - \left(\mu - \frac{\sigma^2}{2} \right) \delta \right)^2}{2\delta\sigma^2} \right)}{\prod_{i=1}^n \frac{e^{-\lambda\delta}}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n! \sqrt{(\delta\sigma^2 + n\sigma_u^2)}} * \exp \left(-\frac{\left(r_i - \left(\mu - \frac{\sigma^2}{2} \right) \delta \right)^2}{2(\delta\sigma^2 + n\sigma_u^2)} \right) \right)}$$

Avec r_i les rendements observés de l'échantillon.

On conclut le test en comparant la valeur $-2\ln(L_{Calculé})$ à la valeur L_{seuil} qui suit une loi de Chi2 à p degrés de liberté ($L_{seuil} \sim \chi^2(p)$), p représentant la différence de paramètres entre les deux modèles, pour un risque α donné :

- $-2\ln(L_{Calculé}) > L_{seuil}$ alors l'hypothèse H_0 est rejetée, le modèle à saut semble meilleur ;
- $-2\ln(L_{Calculé}) \leq L_{seuil}$ alors l'hypothèse H_0 est acceptée, le modèle à saut n'est pas adapté.

b) Test d'adéquation du Chi2

L'objectif de ce test est de pouvoir identifier si les rendements sont perturbés par des sauts et ne suivent donc pas une loi normale comme dans le modèle de Black et Scholes.

Soit X une variable aléatoire, on pose :

$F_E(X)$ la fonction de distribution de X observée;

$F_T(X)$ la fonction de distribution de la loi théorique.

On va tester l'hypothèse H_0 contre l'hypothèse H_1 :

$$\begin{cases} H_0 : F_E(X) = F_T(X) \\ H_1 : F_E(X) \neq F_T(X) \end{cases}$$

Si H_0 est vérifiée, la fonction de distribution est comparable à la fonction de répartition théorique.

La statistique de test, qui peut être calculée une fois les rendements empiriques rangés en K classes, est donnée par :

$$T_N = \sum_{k=1}^K \frac{(n_i - Np_i)^2}{Np_i}$$

Avec K le nombre de classes dans lesquels sont séparés les rendements

N le nombre de données de l'échantillon

n_i le nombre de données dans la classe k , on a : $\sum_{k=1}^K n_k = N$

p_i la probabilité qu'un événement appartienne à la classe i .

On conclut le test en comparant la valeur T_N à la valeur T_{seuil} qui suit une loi de Chi2 à $(K - p - 1)$ degrés de liberté ($T_{seuil} \sim \chi^2(K - p - 1)$), p représentant le nombre de paramètres du modèle testé, pour un risque α donné :

- $T_N > T_{seuil}$ alors l'hypothèse H_0 est rejetée, le modèle ne correspond pas à la distribution empirique ;
- $T_N \leq T_{seuil}$ alors l'hypothèse H_0 est acceptée, le modèle théorique ne peut pas être rejeté.

Preuve :

On classe les rendements empiriques en K classes, et on note n_i le nombre de données dans

la classe k , on a : $\sum_{k=1}^K n_k = N$;

Remarque : Le choix des classes doit respecter le critère de Cochran, elles doivent être assez nombreuses pour ne pas perdre d'information, mais pas trop petites non plus. Concrètement, il faut qu'aucune classe ne soit vide et que 80% d'entre elles contiennent au moins cinq valeurs.

Sous l'hypothèse H_0 , n_i appartient à la classe (k_i, k_{i+1}) , et suit une loi binomiale $B(N, p_i)$, avec $p_i = F(k_{i+1}) - F(k_i)$ la probabilité qu'un événement appartienne à la classe i . On a donc $E[n_i] = Np_i$ et $Var[n_i] = Np_i(1 - p_i)$.

D'après le théorème central limite, sous H_0 :

$$\frac{(n_i - Np_i)}{\sqrt{Np_i}} \xrightarrow{N \rightarrow \infty} N(0, \sqrt{1 - p_i})$$

Propriété

Soit $X_1, X_2, \dots, X_k$ des variables aléatoires suivant une loi normale centrée, réduite, alors :

$Y = \sum_{i=1}^K X_i$ suit une loi de Chi2 à k degrés de liberté.

La statistique de test $T_N = \sum_{k=1}^K \frac{(n_i - Np_i)^2}{Np_i}$, sous l'hypothèse H_0 , converge en loi vers une loi de Chi2 à $K-p-1$ degrés de liberté.

Pour l'application au modèle de Merton, on cherche ensuite à obtenir la fonction de répartition des rendements.

La densité est donnée par :

$$f_r(x) = \frac{e^{-\lambda\delta}}{\sqrt{2\pi}} \left[\sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n! \sqrt{(\sigma^2\delta + n\sigma_u^2)}} * \exp \left(-\frac{\left(x - \left(\mu - \frac{\sigma^2}{2} \right) \delta \right)^2}{2 * (\sigma^2\delta + n\sigma_u^2)} \right) \right) \right]$$

En posant : $\sigma_n^2 = (\sigma^2\delta + n\sigma_u^2)$ et $\mu_n = \left(\mu - \frac{\sigma^2}{2} \right) \delta$

On obtient : $f_r(x) = \frac{e^{-\lambda\delta}}{\sqrt{2\pi}} \left[\sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n! \sigma_n} * \exp \left(-\frac{(x - \mu_n)^2}{2\sigma_n^2} \right) \right) \right]$

La fonction de répartition se déduit de la formule :

$$F_r(x) = \int_{-\infty}^x f_r(t) dt$$

$$F_r(x) = \int_{-\infty}^x \frac{e^{-\lambda\delta}}{\sqrt{2\pi}} \left[\sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n! \sigma_n} * \exp \left(-\frac{(t - \mu_n)^2}{2\sigma_n^2} \right) \right) \right] dt$$

$$F_r(x) = e^{-\lambda\delta} \sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n!} * \int_{-\infty}^x \frac{1}{\sigma_n \sqrt{2\pi}} \exp \left(-\frac{(t - \mu_n)^2}{2\sigma_n^2} \right) dt \right)$$

On reconnaît alors la loi de répartition de la loi normale $N(\mu_n, \sigma_n^2)$:

$$\int_{-\infty}^x \frac{1}{\sigma_n \sqrt{2\pi}} \exp \left(-\frac{(t - \mu_n)^2}{2\sigma_n^2} \right) dt \text{ incluse dans la fonction de répartition calculée.}$$

c) Test d'indépendance des rendements

Comme nous l'avons spécifié dans la présentation du modèle, les rendements doivent être indépendamment et identiquement distribués (iid). Pour s'assurer que ce postulat est vérifié, nous pouvons utiliser le test « up and down », appelé aussi le test des séquences.

Pour réaliser le test, il convient tout d'abord de transformer les rendements en variables dichotomiques qui formeront les séquences¹² constituées de 0 et 1. Le but étant d'analyser les mouvements haussiers ou baissiers, on crée les variables aléatoires Y_i avec $i \in [1, n-1]$, telles que :

$$Y_i = \begin{cases} 1 & \text{si } r_{i+1} - r_i > 0 \\ 0 & \text{si } r_{i+1} - r_i < 0 \end{cases}$$

La réalisation des Y_i permet d'obtenir le nombre de séquences observées, noté R .

On va tester l'hypothèse H_0 contre l'hypothèse H_1 :

$$\begin{cases} H_0 : y_i & iid \\ H_1 : y_i & noniid \end{cases}$$

Si H_0 est vérifiée, les rendements seront indépendants.

Ce test s'applique sur un grand échantillon de données (> 30), grâce à la loi des grands nombres, il est possible de dire que la loi suivie par la variable R est une loi normale.

La statistique de test est :

$$I_{\text{Calculé}} = \frac{R - E(R)}{\sigma(R)}$$

Avec les nombres d'évènements de type 0 et 1 respectivement.

Sous l'hypothèse H_0 , on a :

$$E(R) = \frac{2n-1}{3} \text{ l'espérance de } R$$

$$\sigma^2(R) = \frac{16n-29}{90} \text{ la variance de } R$$

Remarque : pour plus d'information, il est possible de consulter l'article « Probability table for number of runs of signs of first differences in ordered series » de Edgington E.S.

On conclut sur le test en comparant la valeur $I_{\text{Calculé}}$ à la valeur I_{seuil} qui suit une loi de $N(0,1)$, pour un risque α donné :

- $I_{\text{Calculé}} > I_{\text{seuil}}$ alors l'hypothèse de normalité est rejetée, la normalité de l'échantillon est rejetée ;
- $I_{\text{Calculé}} \leq I_{\text{seuil}}$ alors l'hypothèse de normalité ne peut pas être rejetée.

E. Critères de comparaison des modèles

Réaliser un bon calibrage des modèles nécessite souvent de faire plusieurs tests. Une fois que plusieurs modèles répondent relativement correctement en terme d'ajustement et de vérification des hypothèses, comment choisir l'un ou l'autre des modèles ?

Dans la littérature, différents critères ont été proposés pour permettre la comparaison entre les modèles. Les trois critères retenus dans notre étude portent sur la log vraisemblance. Ils

¹² Séquence : suite d'évènement identique, par exemple :

- (0,0,0,1,1,0,1,0) : 5 séquences
- (1,1,1,0,0,0,1,1) : 3 séquences

sont donc applicables que lorsque les estimations seront faites par la méthode du maximum de vraisemblance.

On obtient la log vraisemblance d'un modèle en calculant la valeur pour les paramètres optimums trouvés.

1. Critère Likelihood Ratio Test (LRT)

Le critère LRT « Likelihood Ratio Test » plus connu sous le nom du rapport de vraisemblance, permet de comparer deux modèles, principalement lorsque ceux-ci sont emboîtés (c'est-à-dire que la différence entre les modèles est l'ajout d'un terme explicatif).

On cherche à tester l'hypothèse H_0 contre l'hypothèse H_1 :

$$\begin{cases} H_0 : p_1 = p_2 = \dots = p_p = 0 \\ H_1 : p_1 = p_2 = \dots = p_p \neq 0 \end{cases}$$

Si H_0 est vérifiée, le modèle ayant le moins de variables explicatives sera le meilleur.

La statistique de test étudiée est :

$$LRT_{\text{Calculé}} = -2 * \log\left(\frac{L_1}{L_2}\right)$$
$$LRT_{\text{Calculé}} = -2 * (\log L_1 - \log L_2)$$

Avec L_1 la vraisemblance du modèle ayant le moins de variables explicatives ;

L_2 la vraisemblance du modèle ayant le plus de variables explicatives ;

Sous H_0 , $LRT_{\text{Calculé}}$ suit une loi de Chi2 à p degrés de liberté, p représentant la différence de paramètres entre les deux modèles.

Conclure la comparaison des modèles :

On compare la valeur $LRT_{\text{Calculé}}$ à la valeur LRT_{seuil} qui suit une loi de Chi2 à p degrés de liberté ($LRT_{\text{seuil}} \sim \chi^2(p)$), pour un risque α donné :

- $LRT_{\text{Calculé}} > LRT_{\text{seuil}}$ alors l'hypothèse H_0 est rejetée, le modèle contenant le moins de variables explicatives n'est pas le meilleur ;
- $LRT_{\text{Calculé}} \leq LRT_{\text{seuil}}$ alors l'hypothèse H_0 est acceptée, le modèle avec le moins de variables explicatives est le meilleur.

2. Critère Akaike (AIC / AICc)

Le « Akaike's Information Criterion » développé par Hirotugu Akaike en 1974 est bien souvent utilisé pour comparer la performance relative de deux modèles. Le critère est défini par :

$$AIC = -2 * \log(L) + 2 * k$$

Avec L la log vraisemblance ;

k le nombre de paramètres estimés.

L'intérêt d'utiliser le nombre de paramètres k est de pénaliser les modèles surparamétrés qui obtiennent de ce fait des valeurs de vraisemblance plus importantes.

Ce critère, en l'état, est utilisable uniquement sur de gros échantillons.

Pour les échantillons de taille plus modeste, vérifiant $\frac{n}{k} < 40$, Hurvich et Tsai proposent, en 1989 et 1995, le « AIC corrigé » :

$$AICc = AIC + \frac{2k(k+1)}{n-k-1}$$

Avec L la log vraisemblance ;
 k le nombre de paramètres estimés ;
 n le nombre de données.

Conclure la comparaison des modèles :

Le meilleur des modèles est celui possédant le plus petit AIC (ou AICc).

3. Critère Bayésien (BIC)

Le « Bayesian Information Criterion », communément appelé BIC, a été introduit par Schwartz en 1978.

Le critère pénalisé est proche de l'AIC, il est défini par la formule :

$$BIC = -2 * \log(L) + k * \log(n)$$

Avec L la log vraisemblance ;
 k le nombre de paramètres estimés ;
 n le nombre de données.

Conclure la comparaison des modèles :

Le meilleur des modèles est celui possédant le plus petit BIC.

Remarque : en général, le critère AIC est préféré au BIC lorsqu'il s'agit de trouver un modèle pour réaliser des prévisions (cf. Ripley 2003).

IV. Mise en œuvre du calibrage des classes d'actifs

Cette section se consacre à l'application des méthodes de calibrage, l'objectif étant de trouver pour chacune des classes d'actif le modèle et le calibrage approprié.

Le choix des données est une phase essentielle du calibrage, de ce choix dépendra les résultats, il est donc important de se poser les bonnes questions :

- Quelles données a-t-on à notre disposition ?
- Comment choisir la profondeur d'historique ?
- Comment choisir la granularité ?
- Est-il nécessaire de retravailler les données pour les rendre exploitables ?

Pour toutes ces questions essentielles dans le choix des données, il n'y a pas de réponse standard à apporter. A chacun d'analyser précisément son besoin et les éléments dont il dispose.

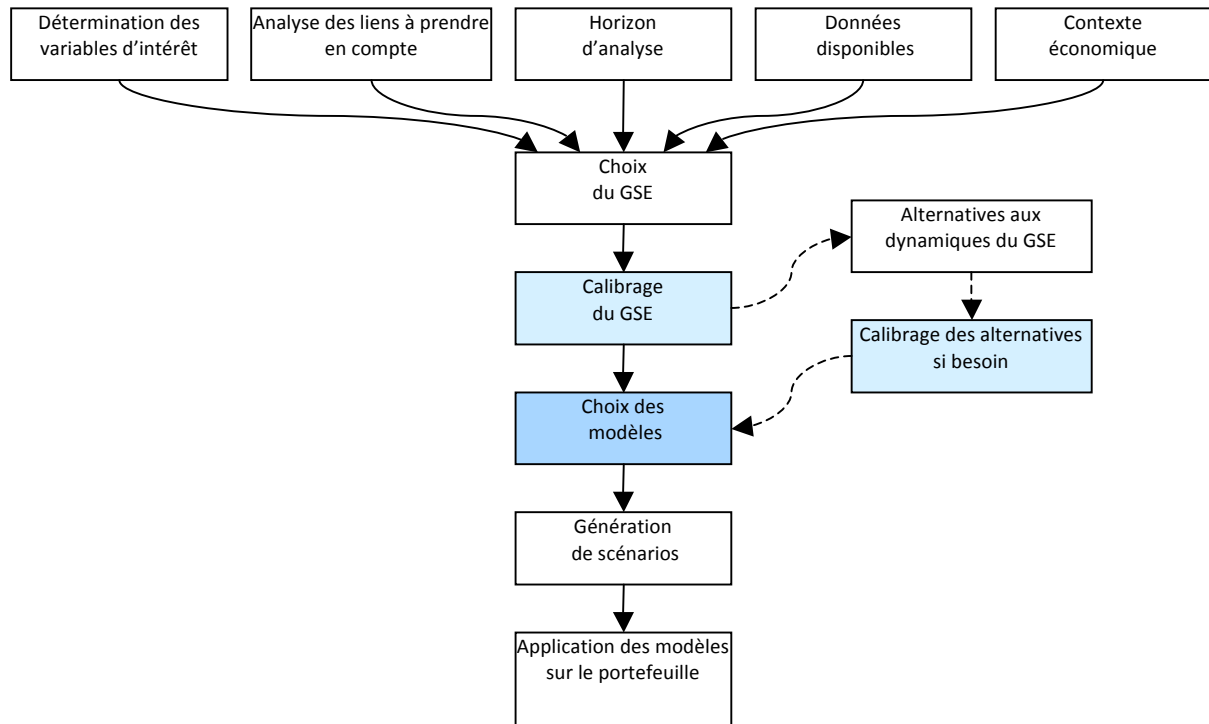
Notre objectif est de réaliser une projection des valeurs d'actifs à long terme (30 ans environ), pour cela, les données de Friggit, souvent citées comme référence conviendront (*cf. §Données de Friggit*).

Usuellement, la profondeur d'historique retenue correspond à la longueur de projection (soit 30 ans ici), nous essayerons de voir si c'est ce qu'il convient de retenir.

La granularité s'impose à nous car les données qui sont annuelles. La projection étant à long terme, il n'est de toute façon pas obligatoire d'avoir des données fines (trimestrielles, mensuelles ou journalières) pour obtenir un modèle satisfaisant.

Par ailleurs, nous avons retenu les données de bourse pour l'action AXA (*cf. §Données de l'action AXA*), dont la granularité est mensuelle, pour étudier les actions de façon plus détaillée. La question est alors de savoir si le même calibrage que pour les données d'action de Friggit conviendra.

La modélisation des données ne peut être faite sur les données brutes, ce sont les rendements qui sont étudiés. Cela s'explique simplement par le fait qu'un accroissement entre deux valeurs n'a en général pas de sens, c'est l'accroissement relatif qui donne une vraie information.

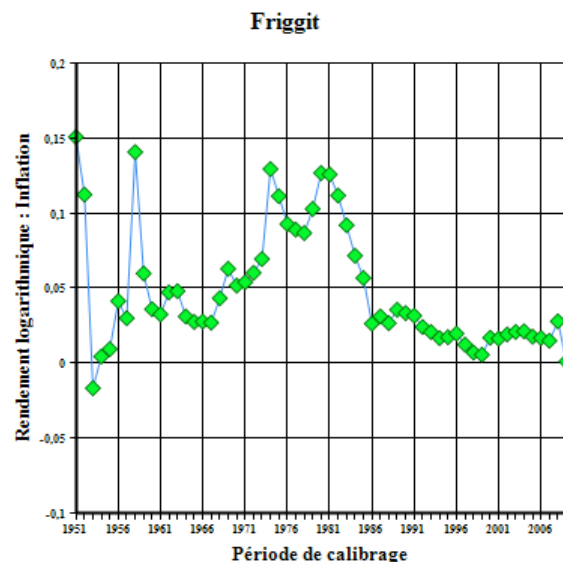
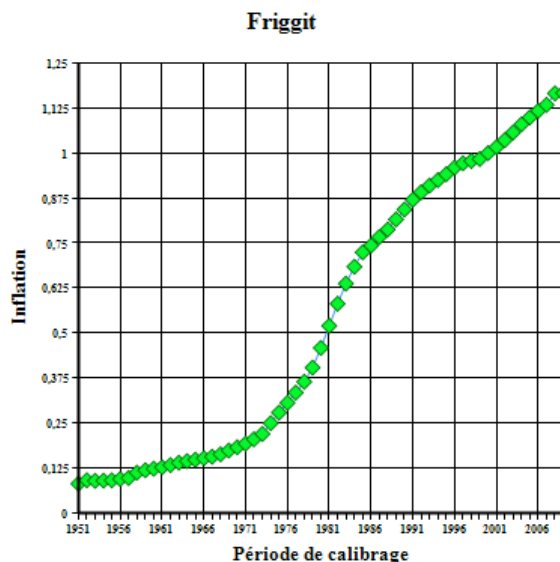


A. L'inflation

1. Les données

Les données de l'inflation de Friggit sont sous leurs formes brutes en montant cumulé. Pour les rendre exploitables nous calculons les rendements logarithmiques de la façon suivante :

$$\text{« Rendement Inflation »}_i = \text{Log Népérien} (\text{Inflation}_i / \text{Inflation}_{i-1})$$



2. Modèle de Vasicek

La modélisation de l'inflation est réalisée à l'aide du processus de Vasicek à un facteur. Le modèle a été sommairement présenté dans les parties précédentes, nous détaillons ici les calculs. La dynamique suivie par l'inflation est :

$$dr_t = (a - br_t)dt + \alpha dB$$

La discrétisation exacte du processus signifie qu'il est possible d'obtenir la solution exacte de l'équation différentielle stochastique, cette solution est donnée par :

$$r_{t_{i+1}} = \exp(-a\delta)r_{t_i} + \frac{b}{a}[1 - \exp(-a\delta)] + \sigma\sqrt{\frac{1 - \exp(-2a\delta)}{2a}}\varepsilon_{t_i}$$

Avec ε_{t_i} une variable aléatoire de loi normale centrée réduite.

Le détail des calculs est donné en annexe (cf. §Modèle de Vasicek à un facteur).

a) Estimation par les moindres carrés ordinaires

Estimation des paramètres

Pour procéder à l'évaluation du modèle de Vasicek, nous utilisons la discrétisation exacte.

En posant : $\alpha = \exp^{-a\delta}$; $\beta = \frac{b}{a}(1 - \exp^{-a\delta})$; $v = \sigma\sqrt{\frac{1 - \exp^{-2a\delta}}{2a}}$

L'équation se simplifie en une équation linéaire :

$$r_{t_{i+1}} = \alpha r_{t_i} + \beta + v\varepsilon_i$$

Cette équation peut se résoudre par la méthode des moindres carrés ordinaires afin d'obtenir les estimateurs de α , β et v .

L'équation de départ ($r_{t_{i+1}} = \alpha r_{t_i} + \beta + v\varepsilon_i$), peut être exprimée sous forme matricielle :

$$y = X\theta + v\varepsilon_i$$

Avec $y = \begin{bmatrix} r_{t_1} \\ r_{t_2} \\ \vdots \\ r_{t_n} \end{bmatrix}$, $X = \begin{bmatrix} 1 & r_{t_0} \\ 1 & r_{t_1} \\ \vdots & \vdots \\ 1 & r_{t_{n-1}} \end{bmatrix}$, $\theta = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$.

Le modèle possède une seule variable explicative r_{t_i} .

Nous cherchons d'abord à estimer θ , son expression par les moindres carrés est :

$$\hat{\theta} = (X'X)^{-1}X'y.$$

A partir de $\hat{\alpha}$ et $\hat{\beta}$, on obtient la valeur estimée des paramètres $\hat{a}$, $\hat{b}$ et $\hat{v}$:

- $\alpha = \exp^{-a\delta} \Rightarrow \hat{a} = -\frac{\ln(\hat{\alpha})}{\delta}$
- $\beta = \frac{b}{a}(1 - \exp^{-a\delta}) \Rightarrow \hat{b} = \frac{\hat{a}\hat{\beta}}{(1 - \exp^{-\hat{a}\delta})} = \frac{-\ln(\hat{\alpha})\hat{\beta}}{\delta(1 - \exp^{\ln(\hat{\alpha})})}$
- $v\varepsilon = \sigma\sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \Rightarrow \hat{v}\hat{\varepsilon} = \sigma\sqrt{\frac{1 - \exp^{-2\hat{a}\delta}}{2\hat{a}}} = \sigma\sqrt{\frac{\delta(1 - \exp^{2\ln(\hat{\alpha})})}{-2\ln(\hat{\alpha})}}$

$$\Rightarrow \hat{\sigma} = \sqrt{\frac{2\hat{a}}{1 - \exp^{-2\hat{a}\delta}} * \frac{\sum_i (\varepsilon_i)^2}{n - 2 - 1}}$$

Matrice de corrélation des paramètres

L'inconvénient majeur de cette méthode, c'est qu'elle ne permet pas d'estimer les valeurs des paramètres a et b de façon directe.

On estime α , β pour obtenir une estimation de a , b pour en déduire σ .

Il n'est donc pas évident d'obtenir les incertitudes et corrélations du modèle car l'estimation n'est pas directe.

Conclusion

Cette méthode n'est pas très adaptée car il est difficile d'obtenir des informations (incertitudes, corrélations) permettant de valider le modèle.

Nous allons donc étudier la méthode des moindres carrés non linéaires qui devrait pallier à l'inconvénient majeur des MCO de ne pas estimer directement les paramètres.

b) Estimation par les moindres carrés non linéaires

Estimation des paramètres

Pour l'application de la méthode des moindres carrés non linéaires au modèle de Vasicek, on pose :

- $\theta = \begin{bmatrix} a \\ b \end{bmatrix}$ le vecteur des paramètres à estimer,
- f une fonction de θ et r_{t_i} : $f(a, b, r_{t_i}) = \exp(-a\delta)r_{t_i} + \frac{b}{a}[1 - \exp(-a\delta)]$,
- $\varepsilon'_i = v\varepsilon_i$, vecteur homoscedastique¹³ de moyenne nulle.

Le modèle s'écrit alors : $r_{t_{i+1}} = f(\theta, r_{t_i}) + \varepsilon'_i$, on se retrouve dans le cadre du modèle non linéaire.

Pour obtenir l'estimation du vecteur de paramètres θ , on utilise l'algorithme de minimisation de Gauss Newton basé sur le gradient.

Le gradient est déterminé par la formule :

$$\nabla SCE(\theta) = \begin{bmatrix} \frac{\partial SCE(\theta)}{\partial a} \\ \frac{\partial SCE(\theta)}{\partial b} \end{bmatrix} = \begin{bmatrix} -2 \sum_{i=0}^{n-1} (r_{t_{i+1}} - f(a, b, r_{t_i})) \frac{\partial f(a, b, r_{t_i})}{\partial a} \\ -2 \sum_{i=0}^{n-1} (r_{t_{i+1}} - f(a, b, r_{t_i})) \frac{\partial f(a, b, r_{t_i})}{\partial b} \end{bmatrix}$$

$$\text{Avec } \frac{\partial f(a, b, r_{t_i})}{\partial a} = -\delta r_{t_i} \exp^{-a\delta} - \frac{b}{a^2} (1 - \exp^{-a\delta}) + \frac{b}{a} \delta \exp^{-a\delta}$$
$$\frac{\partial f(a, b, r_{t_i})}{\partial b} = \frac{1 - \exp^{-a\delta}}{a}$$

Matrice de variance-covariance / corrélation

En utilisant les éléments déjà calculés pour obtenir le gradient, on obtient rapidement la matrice hessienne, qui est donnée par :

¹³ Homoscedasticité : la variance des termes d'erreur est constante pour chaque observation.

$$H(\theta) = \begin{bmatrix} \frac{\delta^2 SCE(\theta)}{\delta a^2} & \frac{\delta^2 SCE(\theta)}{\delta a \delta b} \\ \frac{\delta^2 SCE(\theta)}{\delta b \delta a} & \frac{\delta^2 SCE(\theta)}{\delta b^2} \end{bmatrix}$$

Avec

$$\begin{aligned} & \frac{\delta^2 f(a, b, r_{t_i})}{\delta a^2} = \frac{\delta}{\delta a} \left(\frac{\delta f(a, b, r_{t_i})}{\delta a} + (r_{t_{i+1}} - f(a, b, r_{t_i})) \frac{\delta^2 f(a, b, r_{t_i})}{\delta a^2} - 2 \sum_{i=0}^{n-1} \frac{\delta f(a, b, r_{t_i})}{\delta a} \frac{\delta f(a, b, r_{t_i})}{\delta b} + (r_{t_{i+1}} - f(a, b, r_{t_i})) \frac{\delta^2 f(a, b, r_{t_i})}{\delta a \delta b} \right) \\ & \frac{\delta^2 f(a, b, r_{t_i})}{\delta a \delta b} = \frac{\delta}{\delta a} \left(\frac{\delta f(a, b, r_{t_i})}{\delta b} + (r_{t_{i+1}} - f(a, b, r_{t_i})) \frac{\delta^2 f(a, b, r_{t_i})}{\delta a \delta b} + \frac{b \delta}{a^2} \exp^{-a \delta} \left(\frac{\delta f(a, b, r_{t_i})}{\delta b} + (r_{t_{i+1}} - f(a, b, r_{t_i})) \frac{\delta^2 f(a, b, r_{t_i})}{\delta a \delta b} + \frac{b}{a} \frac{\delta^2 f(a, b, r_{t_i})}{\delta b^2} \right) \right) \\ & = \delta^2 r_{t_i} \exp^{-a \delta} + \frac{2b}{a^3} (1 - \exp^{-a \delta}) - 2 \frac{b \delta}{a^2} \exp^{-a \delta} - \frac{b}{a} \delta^2 \exp^{-a \delta} \\ & \frac{\delta^2 f(a, b, r_{t_i})}{\delta a \delta b} = \frac{\delta^2 f(a, b, r_{t_i})}{\delta b \delta a} = -\frac{1}{a^2} (1 - \exp^{-a \delta}) + \frac{1}{a} \delta \exp^{-a \delta} \\ & \frac{\delta^2 f(a, b, r_{t_i})}{\delta b^2} = 0 \end{aligned}$$

Il est alors possible d'en déduire la matrice de corrélation des paramètres :

$$Corr(\theta) = \begin{pmatrix} 1 & Corr(a, b) \\ Corr(b, a) & 1 \end{pmatrix}$$

En utilisant la matrice de variance-covariance, il est possible de déduire la matrice de corrélation. Les valeurs des variances des paramètres a et b sont données par la diagonale de la matrice de covariance.

Remarque : cette méthode ne permet pas d'obtenir une matrice de corrélation nous donnant des informations sur les trois paramètres estimés : a , b et σ .

c) Estimation par le maximum de vraisemblance

Estimation des paramètres

A partir de l'échantillon à n observations, nous allons chercher à déterminer l'estimation du

paramètre $\theta = \begin{bmatrix} a \\ b \\ \sigma \end{bmatrix}$ optimale.

Étape 1 : Trouver la fonction de densité de la variable aléatoire

Le modèle de Vasicek définit le processus du taux d'intérêt par l'équation :

$$r_{t_{i+1}} = \exp(-a\delta)r_{t_i} + \frac{b}{a}[1 - \exp(-a\delta)] + \sigma \int_{t_i}^{t_{i+1}} \exp(-a(t_{i+1} - s)) dB_s$$

Comme le relatent Brigo et Mercurio dans leur livre, $r_{t_{i+1}}$ est normalement distribué de moyenne et de variance :

$$\begin{aligned} E[r_{t_{i+1}} / r_{t_i}] &= \exp(-a\delta)r_{t_i} + \frac{b}{a}[1 - \exp(-a\delta)] \\ Var[r_{t_{i+1}} / r_{t_i}] &= \sigma^2 \frac{1 - \exp[-2a\delta]}{2a} \end{aligned}$$

Étape 2 : Définir la fonction de log vraisemblance et sa maximisation

La fonction de vraisemblance est donnée par :

$$L(\theta) = \prod_{i=0}^{n-1} \frac{1}{\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{r_{t_{i+1}} - \left(\exp(-a\delta) r_{t_i} + \frac{b}{a} [1 - \exp(-a\delta)] \right)}{\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}}} \right)^2 \right)$$

La fonction de log vraisemblance devient alors :

$$\Lambda = -n \log \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \sqrt{2\pi} \right) - \frac{1}{2 \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \right)^2} \sum_{i=0}^{n-1} \left(r_{t_{i+1}} - \exp^{-2a\delta} r_{t_i} - \frac{b}{a} (1 - \exp^{-2a\delta}) \right)^2$$

où $\Lambda = \sum_{i=0}^{n-1} \Lambda_{i+1}$ avec

$$\Lambda_{i+1} = -\log \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \sqrt{2\pi} \right) - \frac{1}{2 \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \sqrt{2\pi} \right)^2} \left(r_{t_{i+1}} - \exp^{-2a\delta} r_{t_i} - \frac{b}{a} (1 - \exp^{-2a\delta}) \right)^2$$

Pour tenter de résoudre le problème, on simplifie la fonction de vraisemblance, en posant :

$$\alpha = \exp^{-a\delta}, \quad \beta = \frac{b}{a} (1 - \alpha) \quad \text{et} \quad v = \sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}}$$

La fonction de vraisemblance est :

$$L(\theta) = \prod_{i=0}^{n-1} \frac{1}{v \sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{r_{t_{i+1}} - (\alpha r_{t_i} + \beta [1 - \alpha])}{v} \right)^2 \right)$$

La fonction de log vraisemblance devient alors :

$$\Lambda = -n \log(v \sqrt{2\pi}) - \frac{1}{2v^2} \sum_{i=0}^{n-1} (r_{t_{i+1}} - \alpha r_{t_i} - \beta(1 - \alpha))^2$$

$$\text{où } \Lambda = \sum_{i=0}^{n-1} \Lambda_{i+1} \quad \text{avec} \quad \Lambda_{i+1} = -\log(v \sqrt{2\pi}) - \frac{1}{2(v \sqrt{2\pi})^2} (r_{t_{i+1}} - \alpha r_{t_i} - \beta(1 - \alpha))^2$$

Le gradient des résidus :

$$\begin{aligned} \frac{\partial \varepsilon'_{i+1}}{\partial \alpha} &= -r_{t_i} + \beta \\ \frac{\partial \varepsilon'_{i+1}}{\partial \beta} &= -(1 - \alpha) \\ \frac{\partial \varepsilon'_{i+1}}{\partial v} &= 0 \end{aligned}$$

Le gradient de log vraisemblance :

$$\begin{aligned}\frac{\partial \Lambda_{i+1}}{\partial \alpha} &= -\frac{\varepsilon'_{i+1} \frac{\partial \varepsilon'_{i+1}}{\partial \alpha}}{v^2} \\ \frac{\partial \Lambda_{i+1}}{\partial \beta} &= -\frac{\varepsilon'_{i+1} \frac{\partial \varepsilon'_{i+1}}{\partial \beta}}{v^2} \\ \frac{\partial \Lambda_{i+1}}{\partial v} &= -\frac{1}{v} + \frac{\varepsilon'^2_{i+1}}{v^3}\end{aligned}$$

La vraisemblance sera maximisée pour les paramètres vérifiant le système d'équations :

$$\left\{ \begin{aligned} \sum_{i=0}^{n-1} \left(\frac{\partial \Lambda_{i+1}}{\partial \alpha} \right) &= -\sum_{i=0}^{n-1} \left(\frac{\varepsilon'_{i+1} \frac{\partial \varepsilon'_{i+1}}{\partial \alpha}}{v^2} \right) = 0 \\ \sum_{i=0}^{n-1} \left(\frac{\partial \Lambda_{i+1}}{\partial \beta} \right) &= -\sum_{i=0}^{n-1} \left(\frac{\varepsilon'_{i+1} \frac{\partial \varepsilon'_{i+1}}{\partial \beta}}{v^2} \right) = 0 \\ \sum_{i=0}^{n-1} \left(\frac{\partial \Lambda_{i+1}}{\partial v} \right) &= \sum_{i=0}^{n-1} \left(-\frac{1}{v} + \frac{\varepsilon'^2_{i+1}}{v^3} \right) = 0 \end{aligned} \right.$$

En observant les différentes équations, la première constatation est que v est non nul. On déduit sa valeur de la troisième équation du système :

$$\hat{v}^2 = \frac{\sum_{i=0}^{n-1} (\varepsilon'^2_{i+1})}{n}$$

De la seconde équation, on déduit l'expression de β :

$$\hat{\beta} = \frac{\sum_{i=0}^{n-1} (r_{i+1}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n(1 - \alpha)}$$

Enfin l'expression de α est déduite de la première équation :

$$\hat{\alpha} = \frac{n \sum_{i=0}^{n-1} (r_{i+1} r_{t_i}) - \sum_{i=0}^{n-1} (r_{t_i}) \sum_{i=0}^{n-1} (r_{i+1})}{n \sum_{i=0}^{n-1} (r_{t_i}^2) - \sum_{i=0}^{n-1} (r_{t_i})^2}$$

A partir de $\hat{\alpha}$ et $\hat{\beta}$, on obtient la valeur estimée des paramètres $\hat{a}$, $\hat{b}$ et $\hat{v}$:

- $\alpha = \exp^{-a\delta} \Rightarrow \hat{a} = -\frac{\ln(\hat{\alpha})}{\delta}$
- $\beta = \frac{b}{a}(1 - \alpha) \Rightarrow \hat{b} = \frac{\hat{a}\hat{\beta}}{1 - \hat{\alpha}}$

$$\begin{aligned} \bullet \quad v &= \sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \Rightarrow \hat{v} = \sigma \sqrt{\frac{1 - \exp^{-2\hat{a}\delta}}{2\hat{a}}} = \sigma \sqrt{\frac{\delta(1 - \exp^{2\ln(\hat{\alpha})})}{-2\ln(\hat{\alpha})}} \\ &\Rightarrow \hat{\sigma} = \sqrt{\frac{2\hat{a}}{1 - \exp^{-2\hat{a}\delta}} * \frac{\sum_i (\varepsilon_i)^2}{n}} \end{aligned}$$

Dans le cadre du maximum de vraisemblance, il n'y a pas de degré de liberté à retirer au n .
Remarque : l'expression des paramètres est équivalente à celle de la méthode des MCO.

La moyenne de la variable r_{t_i} , noté $\bar{r}_{t_i}$, est : $\frac{\sum_{i=0}^{n-1} (r_{t_i})}{n}$.

$$\begin{aligned} \hat{\alpha} &= \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) - \frac{\sum_{i=0}^{n-1} (r_{t_i}) \sum_{i=0}^{n-1} (r_{t_{i+1}})}{n}}{\sum_{i=0}^{n-1} (r_{t_i}^2) - \frac{\sum_{i=0}^{n-1} (r_{t_i})^2}{n}} \quad \text{d'où } \hat{\alpha} = \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) - n \bar{r}_{t_{i+1}} \bar{r}_{t_i}}{\sum_{i=0}^{n-1} (r_{t_i}^2) - n \bar{r}_{t_i}^2} ; \\ \hat{\beta} &= \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \hat{\alpha} \sum_{i=0}^{n-1} (r_{t_i})}{n(1 - \hat{\alpha})} \quad \text{d'où } (1 - \hat{\alpha}) \hat{\beta} = \bar{r}_{t_{i+1}} - \hat{\alpha} \bar{r}_{t_i} ; \end{aligned}$$

Pour faciliter les démarches de calcul, il est opportun d'utiliser les valeurs des estimations de $\hat{\alpha}$, $\hat{\beta}$ de la méthode des moindres carrés ordinaires. Ensuite, on déduira $\hat{v}$, $\hat{a}$ et $\hat{b}$.

Matrice de variance-covariance / corrélation

Comme pour l'estimation des paramètres du processus, le calcul s'effectue en deux temps, pour les résidus puis pour la log vraisemblance.

Gradient des résidus : $\varepsilon'_{i+1} = r_{t_{i+1}} - r_{t_i} \exp^{-a\delta} - \frac{b}{a} (1 - \exp^{-a\delta})$

$$\begin{aligned} \frac{\partial \varepsilon'_{i+1}}{\partial a} &= \delta \exp^{-a\delta} r_{t_i} + \frac{b}{a^2} (1 - \exp^{-a\delta}) - \frac{\delta b}{a} \exp^{-a\delta} \\ \frac{\partial \varepsilon'_{i+1}}{\partial b} &= -\frac{1}{a} (1 - \exp^{-a\delta}) \\ \frac{\partial \varepsilon'_{i+1}}{\partial \sigma} &= 0 \end{aligned}$$

Les éléments de la **Hessienne des résidus** :

$$\begin{aligned} \frac{\partial^2 \varepsilon'_{i+1}}{\partial a^2} &= -\delta^2 \exp^{-a\delta} r_{t_i} - \frac{2b}{a^3} (1 - \exp^{-a\delta}) + \frac{2\delta b}{a^2} \exp^{-a\delta} + \frac{\delta^2 b}{a} \exp^{-a\delta} ; \\ \frac{\partial^2 \varepsilon'_{i+1}}{\partial a \partial b} &= \frac{\partial^2 \varepsilon'_{i+1}}{\partial b \partial a} = \frac{(1 - \exp^{-a\delta})}{a^2} - \frac{\delta}{a} \exp^{-a\delta} ; \\ \frac{\partial^2 \varepsilon'_{i+1}}{\partial a \partial \sigma} &= \frac{\partial^2 \varepsilon'_{i+1}}{\partial \sigma \partial a} = 0 ; \end{aligned}$$

$$\begin{aligned}\frac{\partial^2 \varepsilon'_{i+1}}{\partial b^2} &= 0 ; \\ \frac{\partial^2 \varepsilon'_{i+1}}{\partial b \partial \sigma} &= \frac{\partial^2 \varepsilon'_{i+1}}{\partial \sigma \partial b} = 0 ; \\ \frac{\partial^2 \varepsilon'_{i+1}}{\partial \sigma^2} &= 0\end{aligned}$$

Dérivées de la fonction log vraisemblance :

$$\begin{aligned}\frac{\partial \Lambda_{i+1}}{\partial a} &= -\frac{\delta \exp^{-2a\delta}}{1 - \exp^{-2a\delta}} + \frac{0.5}{a} - \frac{\varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} + \frac{2a\delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} - \frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} \\ \frac{\partial \Lambda_{i+1}}{\partial b} &= -\frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial b} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} \\ \frac{\partial \Lambda_{i+1}}{\partial \sigma} &= -\frac{1}{\sigma} + \frac{2a \varepsilon'^2_{i+1}}{\sigma^3 (1 - \exp^{-2a\delta})}\end{aligned}$$

Les éléments de la Hessienne pour la log vraisemblance :

$$\begin{aligned}\frac{\partial^2 \Lambda_{i+1}}{\partial a^2} &= \frac{2\delta^2 \exp^{-2a\delta}}{(1 - \exp^{-2a\delta})} + \frac{2\delta^2 \exp^{-4a\delta}}{(1 - \exp^{-2a\delta})^2} - \frac{0.5}{a^2} - \frac{2 \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} + \frac{2\delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} \\ &+ \frac{2\delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} - \frac{4a\delta^2 \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} + \frac{4a\delta \exp^{-2a\delta} \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} - \frac{8a\delta^2 \exp^{-4a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^3} \\ &- \frac{2 \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} - \frac{2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial a^2} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} - \frac{2a \left(\frac{\partial \varepsilon'_{i+1}}{\partial a} \right)^2}{\sigma^2 (1 - \exp^{-2a\delta})} + \frac{4a\delta \exp^{-2a\delta} \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} \\ \frac{\partial^2 \Lambda_{i+1}}{\partial a \partial b} &= \frac{\partial^2 \Lambda_{i+1}}{\partial b \partial a} = -\frac{2 \frac{\partial \varepsilon'_{i+1}}{\partial b} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} + \frac{4a\delta \exp^{-2a\delta} \frac{\partial \varepsilon'_{i+1}}{\partial b} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} - \frac{2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial a \partial b} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} - \frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \frac{\partial \varepsilon'_{i+1}}{\partial b}}{\sigma^2 (1 - \exp^{-2a\delta})} \\ \frac{\partial^2 \Lambda_{i+1}}{\partial a \partial \sigma} &= \frac{2 \varepsilon'^2_{i+1}}{\sigma^3 (1 - \exp^{-2a\delta})} - \frac{4a\delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^3 (1 - \exp^{-2a\delta})^2} + \frac{4a \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^3 (1 - \exp^{-2a\delta})} \\ \frac{\partial^2 \Lambda_{i+1}}{\partial b^2} &= -\frac{2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial b^2} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} - \frac{2a \left(\frac{\partial \varepsilon'_{i+1}}{\partial b} \right)^2}{\sigma^2 (1 - \exp^{-2a\delta})} \\ \frac{\partial^2 \Lambda_{i+1}}{\partial b \partial \sigma} &= \frac{\partial^2 \Lambda_{i+1}}{\partial \sigma \partial b} = \frac{4a \frac{\partial \varepsilon'_{i+1}}{\partial b} \varepsilon'_{i+1}}{\sigma^3 (1 - \exp^{-2a\delta})}\end{aligned}$$

$$\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma^2} = \frac{1}{\sigma^2} - \frac{6a\varepsilon_{i+1}^2}{\sigma^4(1 - \exp^{-2a\delta})}$$

La matrice Hessienne est alors :

Il ne faut pas oublier que $\Lambda = \sum_{i=0}^{n-1} \Lambda_{i+1}$, la matrice Hessienne est :

$$H_{\Lambda}(\theta) = \begin{pmatrix} \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial a^2} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial a \partial b} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial a \partial \sigma} \right) \\ \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial b \partial a} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial b^2} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial b \partial \sigma} \right) \\ \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma \partial a} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma \partial b} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma^2} \right) \end{pmatrix}$$

Les matrices de variance-covariance, corrélation et les écarts-types se déduisent des formules :

$$\begin{aligned} \text{Cov}(\theta) &= -H(\theta)^{-1} \\ \text{Corr}(X_i, X_j) &= \frac{\text{Cov}(X_i, X_j)}{\sqrt{\text{Var}(X_i)} \sqrt{\text{Var}(X_j)}} \end{aligned}$$

3. Résultat du calibrage

a) Méthode d'estimation retenue

Les trois méthodes d'estimation conduisent aux mêmes valeurs estimées des paramètres pour le processus de Vasicek. Ce qui distingue les modèles c'est la facilité à avoir de l'information sur la qualité d'ajustement du modèle.

Le modèle des MCO ne permet pas d'avoir la matrice de corrélation, c'est pourquoi il était nécessaire d'utiliser d'autres modèles.

Le modèle des moindres carrés non linéaires, quant à lui, permet d'avoir directement la valeur des paramètres et d'en déduire la matrice de corrélation mais sans le paramètre correspondant à la volatilité. Le modèle appliqué au processus de Vasicek donne des équations qu'il est possible de résoudre numériquement, ce qui n'est pas toujours le cas.

L'alternative consistant à utiliser l'estimation par le maximum de vraisemblance, qui donne les paramètres et l'incertitude les concernant, sera privilégiée pour la suite du document.

b) 1^{er} test : 31 données (1978-2009)

Nous souhaitons réaliser une projection sur 30 ans, nous retenons donc un historique équivalent, soit 31 données, ce qui correspond à 30 rendements. Les résultats obtenus sont présentés ci-après.

Valeur des paramètres estimés

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Force de Rappel	0,0933	0,0604	0,1335	La force de Rappel n'est pas significative
Moyenne	0,0001	0,003	0,9828	La moyenne n'est pas significative
Volatilité	0,0112	0,0015	<0.00001	La volatilité est significative

Qualité d'ajustement du modèle

L'analyse de la variance :

Source	Ddl	SCE	CM	F	Prob>F
Modèle	1	0,031325	0,031325	255,785	0
Erreur	28	0,003429	0,000122		
Total	29	0,034754			

Le coefficient de détermination ajusté $R^2_{ajusté}$ est de 0,8978, le pouvoir explicatif du modèle est satisfaisant.

La matrice de corrélation entre les estimateurs :

	Force de Rappel	Moyenne	Volatilité
Force de Rappel	1	0,7381	0,2211
Moyenne	0,7381	1	0,1632
Volatilité	0,2211	0,1632	1

Test d'hypothèse du modèle

Test	Statistique	Probabilité	Interprétation
Moyenne = 0	0	1	La moyenne des résidus est considérée nulle
Variance = 1	30	0,414	La variance des résidus est considérée valoir 1
Shapiro-Wilk	0,9167	0,0221	Hypothèse de normalité des résidus non vérifiée
Breush-Godfrey	0,2234	0,6364	Absence d'autocorrélation d'ordre 1
Arch(1) LM	0,375	0,5403	Absence d'hétéroscédaticité d'ordre 1

c) 2^{ème} test : toute la série de données (1951-2009)

Le premier calibrage donnant des résultats insatisfaisants concernant la significativité des paramètres, un second test est réalisé en prenant toute la série des données. Les résultats obtenus sont présentés ci-après.

Valeur des paramètres estimés

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Force de Rappel	0,3825	0,1277	0,0041	La force de rappel est significative
Moyenne	0,0155	0,0073	0,0385	La moyenne est significative
Volatilité	0,0313	0,0034	<0.00001	La volatilité est significative

Qualité d'ajustement du modèle

L'analyse de la variance :

Source	Ddl	SCE	CM	F	Prob>F
Modèle	1	0,042033	0,042033	59,2	0
Erreur	56	0,039761	0,00071		
Total	57	0,081794			

Le coefficient de détermination ajusté $R^2_{ajusté}$ est de 0,5052, le pouvoir explicatif du modèle n'est pas satisfaisant.

La matrice de corrélation entre les estimateurs :

	Force de Rappel	Moyenne	Volatilité
Force de Rappel	1	0,8242	0,5151
Moyenne	0,8242	1	0,4245
Volatilité	0,5151	0,4245	1

Test d'hypothèse du modèle

Les tests sont réalisés pour un risque d'erreur $\alpha = 5\%$.

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Moyenne = 0	0	1	La moyenne des résidus est considérée nulle
Variance = 1	58	0,4382	La variance des résidus est considérée valoir 1
Shapiro-Wilk	0,7873	0	Hypothèse de normalité des résidus non vérifiée
Breusch-Godfrey	0,0402	0,841	Absence d'autocorrélation d'ordre 1
Arch(1) LM	0,0593	0,8076	Absence d'hétéroscédacité d'ordre 1

d) Conclusion

La significativité des paramètres est admise lorsque l'on utilise toute la série des données pour calibrer le modèle. En revanche, la conclusion sur la qualité d'ajustement des modèles est inverse.

Dans les deux cas, la corrélation entre les paramètres estimés est très importante.

En prenant sur toute la série, le R^2 ajusté n'est pas suffisant pour considérer le modèle comme significatif, il aurait été souhaitable d'avoir un coefficient plus proche de 1.

Il faut cependant avoir à l'esprit que l'étude des processus est particulier, le modèle explique une trajectoire parmi tant d'autres. Il n'y a pas de « solution idéale » à trouver. Les indicateurs habituels de validation des modèles ne sont donc pas toujours totalement adaptés.

Remarque : dans le cas du calibrage d'un modèle classique, une valeur aussi faible du R^2 ajusté aurait conduit au rejet du modèle.

Enfin concernant les tests d'hypothèses, le vecteur de résidus est identique pour les trois méthodes d'estimation, les tests d'hypothèses auraient eu les mêmes résultats quelle que soit la méthode d'estimation retenue.

Remarque : Attention cependant à ne pas utiliser les vecteurs de résidus « bruts » des estimations, il faut supprimer la constante des résidus pour réaliser les tests. En pratique, il n'est pas nécessaire de le faire car la constante n'a pas d'impact sauf pour le test de la variance. Les résidus à tester sont obtenus par :

$$residu_{t_{i+1}} = (r_{t_{i+1}} - r_{t_i} \exp(-a\delta) - \frac{b}{a}(1 - \exp(-a\delta)))$$

$$residu_{t_{i+1}}^{standard} = residu_{t_{i+1}} / (\sigma \sqrt{\frac{1 - \exp(-2a\delta)}{2a}})$$

Un seul test d'hypothèse n'est pas vérifié : celui de la normalité des résidus.

Le calibrage qu'il est préférable de retenir pour l'inflation est celui basé sur toute la série de données, malgré une qualité d'ajustement des données faible.

Pour nous conforter dans notre choix, utilisons les critères de sélection des modèles. Le test du rapport de vraisemblance n'est pas applicable car nous utilisons le même modèle, de même que l'AIC en raison du petit échantillon de données. Voici les résultats obtenus pour l'AICc et le BIC :

	Test 1	Test 2
N	31	59
K	3	3
Log vraisemblance	93,5819	128,9755
Critère AICc	2,9465	2,2153
Critère BIC	-0,9597	-0,6793

Le critère AICc confirme bien qu'il est préférable d'utiliser le calibrage sur l'ensemble de la série.

Validation rétrospective (Backtesting)

Cette méthode consiste, à partir du modèle retenu, à simuler une partie de l'échantillon afin de comparer les résultats simulés aux résultats réels. Le backtesting est souvent utilisé en finance, il permet de valider la méthode à retenir pour le calcul de la VaR (Value-at-Risk¹⁴) des portefeuilles étudiés, la VaR pouvant être obtenue par différents calculs. L'article de Hurlin et Tokpavi propose une « évaluation des procédures de backtesting ».

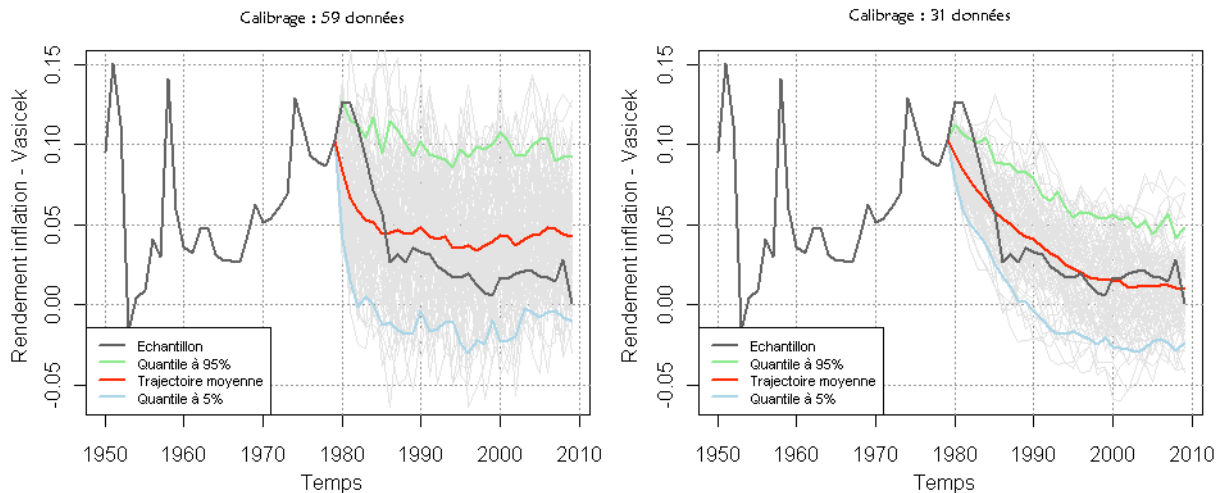
En réalisant le backtesting, la bonne méthode de VaR doit vérifier que les pertes effectives comparées à la VaR ne sont pas plus importantes en fréquence que ce qui est étudié (par exemple : pour la VaR à 99%, dans la réalité seule une donnée sur 100 doit dépasser la valeur de la VaR estimée, s'il y en a plus c'est que la méthode n'est pas la bonne).

Dans notre cas, si le modèle est adapté aux données, on s'attend à ce que la moyenne des simulations soit proche en valeur et structure des données réellement observées. Pour autant il ne faut pas tirer de conclusions hâtives, cela peut signifier que le modèle reproduit bien les données qu'il a calibré, mais n'oublions pas que l'objectif est de réaliser des projections sur les années non connues.

De plus dans le cadre de l'utilisation pour le choix de la VaR en finance, l'utilisation est journalière, dans notre cas la projection est à long terme, la forme du modèle à plus de chance d'être modifiée. Nous utilisons ce procédé à titre indicatif.

¹⁴ Value-at-Risk : dans le cas de l'estimation de la valeur d'un portefeuille de titres, cela correspond au montant de perte qui devrait être dépassé avec une probabilité donnée pour un horizon de temps donné.

Backtesting inflation - Vasicek (Un échantillon de 100 trajectoires)



On s'aperçoit ici que pour le modèle retenu, l'inflation est légèrement surestimée car la courbe rouge est pratiquement toujours au dessus de la courbe des observations (graphique de gauche). Ce qui est plus surprenant sur le second graphique, c'est que les données simulées semblent en moyenne plus proches de la réalité. Le calibrage sur la série complète reste à privilégier.

B. Les taux d'intérêts (court et long)

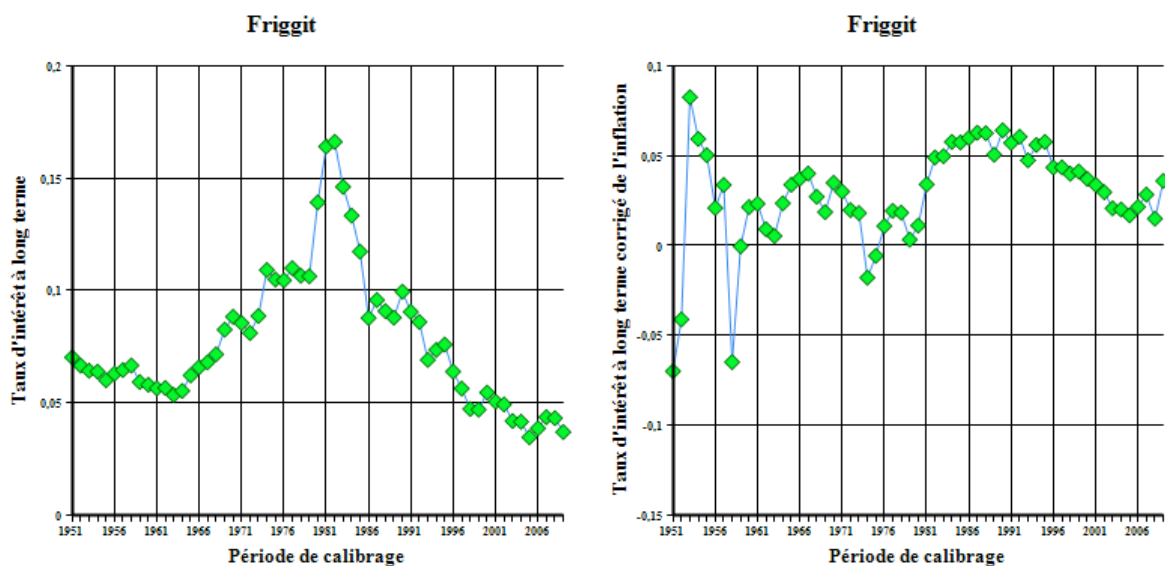
1. Les données

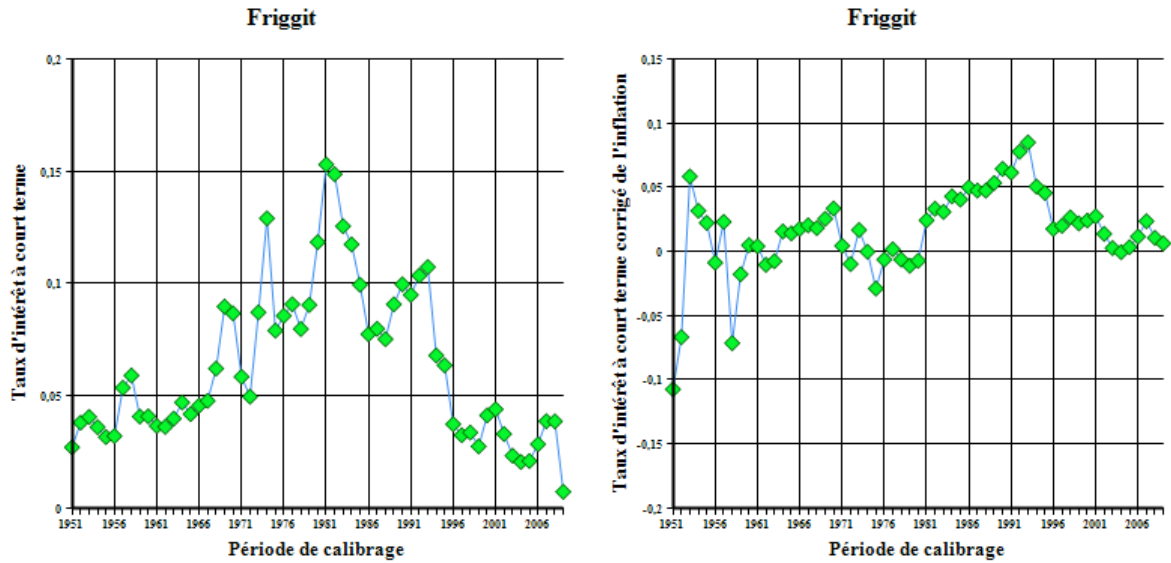
Pour que les données soient exploitables pour une modélisation, l'étude nous oblige à travailler sur les taux corrigés des rendements de l'inflation, pour ne pas biaiser les estimations.

Les taux corrigés suivent alors :

$$\text{« Taux long corrigés »}_i = (1 + \text{« Taux long »}_i) / (1 + \text{« Rendement Inflation »}_i) - 1$$

$$\text{« Taux court corrigés »}_i = (1 + \text{« Taux court »}_i) / (1 + \text{« Rendement Inflation »}_i) - 1$$





2. Modèle de Vasicek à deux facteurs

La dynamique

Le modèle a été sommairement présenté dans les parties précédentes, nous détaillons ici les calculs si nécessaires.

Pour rappel, les dynamiques suivies par le taux court et le taux long sous la probabilité risque neutre, sont :

$$\begin{aligned} dr_t &= a_r(l_t - r_t)dt + \sigma_r dB_{r,t} \\ dl_t &= a_l(\mu_l - l_t)dt + \sigma_l dB_{l,t} \\ dB_r dB_l &= \rho dt \end{aligned}$$

Avec :

- a_r, a_l vitesses de retour à la moyenne ;
- μ_l moyenne à long terme du taux d'intérêt réel long terme ;
- σ_r, σ_l volatilités des taux ;
- $B_{r,t}, B_{l,t}$ mouvements browniens standards indépendant permettant d'introduire l'aléa dans les simulations ;
- ρ corrélation des résidus.

La discrétisation exacte

La discrétisation exacte de chacune des équations se retrouve de la même façon que pour la discrétisation de Vasicek à un facteur (cf. §discrétisation exacte).

On obtient alors les équations :

$$\begin{aligned} r_{t_{i+1}} &= \exp(-a_r \delta) r_{t_i} + l_{t_i} [1 - \exp(-a_r \delta)] + \sigma_r \sqrt{\frac{1 - \exp[-2a_r \delta]}{2a_r}} \varepsilon_{r,t_i} \\ l_{t_{i+1}} &= \exp(-a_l \delta) l_{t_i} + \mu_l [1 - \exp(-a_l \delta)] + \sigma_l \sqrt{\frac{1 - \exp[-2a_l \delta]}{2a_l}} \varepsilon_{l,t_i} \end{aligned}$$

Avec $\varepsilon_{r,t_i}, \varepsilon_{l,t_i}$ des variables aléatoires indépendantes de lois normales centrées réduites.

Le calibrage de ce modèle sera un peu différent de celui du modèle de Vasicek à un facteur car il faut prendre en compte les deux dynamiques.

Il est probable que la volatilité du taux court soit plus forte en raison du double effet des variables aléatoires ε_{r,t_i} , et de ε_{l,t_i} pour le taux long.

a) Estimation par les moindres carrés non linéaires

Estimation des paramètres

Pour procéder à l'évaluation du modèle de Vasicek à deux facteurs, nous utilisons les discrétisations exactes obtenues précédemment.

Les paramètres des deux dynamiques sont indépendants, les paramètres des équations vont pouvoir être estimés séparément. Pour cela, nous appliquerons la méthode des doubles moindres carrés, en procédant en trois étapes, à savoir :

- Estimation des paramètres de la dynamique du taux long ;
- Estimation du taux long $\hat{l}_{t_{i+1}}$;
- Estimation des paramètres de la dynamique du taux court, à l'aide de $\hat{l}_{t_{i+1}}$.

(1) Estimation des paramètres du taux long

L'estimation de ces paramètres se fait de la même façon que pour le modèle Vasicek à un facteur, on utilise ici la méthode des moindres carrés non linéaires. Reportez vous à la partie calibrage de l'inflation pour le détail des calculs du taux long.

Pour estimer le vecteur des paramètres θ_l , on minimise :

$$\text{Min}_{\theta} \text{SCE}(\theta_l) = \text{Min}_{\theta} \sum_{t=1}^T (l_{t_{i+1}} - f(\theta_l, l_{t_i}))^2 = l(\theta_l)'l(\theta_l)$$

avec $f(\theta_l, l_{t_i}) = l_{t_i} \exp(-a_l \delta) + \mu_l [1 - \exp(-a_l \delta)]$

$$\theta_l = \begin{bmatrix} a_l \\ \mu_l \end{bmatrix} \text{ la valeur estimée des paramètres ;}$$

$l(\theta_l)$ la fonction des résidus.

(2) Estimation du taux long $\hat{l}_{t_{i+1}}$

L'estimation des paramètres de la dynamique du taux court va s'appuyer sur les valeurs estimées du taux long. Pour cela on calcule le $\hat{l}_{t_{i+1}}$ de la façon suivante :

$$\hat{l}_{t_{i+1}} = \exp(-a_l \delta) l_{t_i} + \mu_l [1 - \exp(-a_l \delta)]$$

(3) Estimation des paramètres du taux court

Pour estimer les paramètres de la dynamique du taux court terme, il faut tenir compte de l'évaluation du taux de long terme. La discrétisation du taux court terme est alors :

$$r_{t_{i+1}} = \exp(-a_r \delta) r_{t_i} + \hat{l}_{t_i} [1 - \exp(-a_r \delta)] + \sigma_r \sqrt{\frac{1 - \exp[-2a_r \delta]}{2a_r}} \varepsilon_{r,t_i}$$

Avec $\varepsilon_{r,t_i} \sim N(0,1)$.

Remarque : dans cette dynamique, il y a un paramètre de moins à estimer que pour le taux long.

Puis on procède de la même façon que pour le taux long. Pour estimer le vecteur des paramètres θ_r , on minimise :

$$\text{Min}_{\theta} \text{SCE}(\theta_r) = \text{Min}_{\theta} \sum_{t=1}^T (r_{t+1} - f(\theta_r, r_t))^2 = r(\theta_r)'r(\theta_r)$$

avec $f(\theta_r, r_t) = r_t \exp(-a_r \delta) + \hat{l}_t [1 - \exp(-a_r \delta)]$

$\theta_r = \begin{bmatrix} a_r \\ \sigma_r \end{bmatrix}$ la valeur estimée des paramètres ;

$r(\theta_r)$ la fonction des résidus.

$$\nabla \text{SCE}(\theta_r) = -2 \sum_{t=0}^{n-1} (r_{t+1} - f(\theta_r, r_t)) \frac{\partial f(\theta_r, r_t)}{\partial a_r} = 0$$

On applique aussi le même changement de variable, à savoir $\alpha_r = \exp(-a_r \delta)$, ce qui donne :

$$\nabla \text{SCE}(\theta_r) = -2 \sum_{t=0}^{n-1} (r_{t+1} - r_t \alpha_r - \hat{l}_t (1 - \alpha_r)) (-r_t \delta \alpha_r + \hat{l}_t \delta \alpha_r) = 0$$

Il résulte que :

$$\hat{\alpha}_r = \frac{\sum_{t=0}^{n-1} (r_{t+1} r_t) - \sum_{t=0}^{n-1} (r_{t+1} \hat{l}_t) - \sum_{t=0}^{n-1} (r_t \hat{l}_t) + \sum_{t=0}^{n-1} (\hat{l}_t^2)}{\sum_{t=0}^{n-1} (r_t - \hat{l}_t)^2}$$

L'estimation de la volatilité v_r est :

$$v_r = \sigma_r \sqrt{\frac{1 - \exp^{-2a_r \delta}}{2a_r}} \Rightarrow \hat{v}_r \hat{\varepsilon} = \sigma_r \sqrt{\frac{1 - \exp^{-2\hat{a}_r \delta}}{2\hat{a}_r}}$$

$$\Rightarrow \hat{\sigma}_r = \sqrt{\frac{2\hat{a}_r}{1 - \exp^{-2\hat{a}_r \delta}} * \frac{\sum_i (\hat{r}_{i+1})^2}{n}}$$

La matrice Hessienne est donnée par :

$$H(\theta_r) = \left[\frac{\partial^2 \text{SCE}(\theta_r)}{\partial a_r^2} \right] = -2 \sum_{t=0}^{n-1} - \frac{\partial f(a_r, r_t)}{\partial a_r} \frac{\partial f(a_r, r_t)}{\partial a_r} + (r_{t+1} - f(a_r, r_t)) \frac{\partial^2 f(a_r, r_t)}{\partial a_r^2}$$

Avec $\frac{\partial^2 f(a_r, r_t)}{\partial a_r^2} = r_t \delta^2 \exp(-a_r \delta) - \hat{l}_t \delta^2 \exp(-a_r \delta)$

(4) Corrélation entre les résidus des dynamiques

Outre la corrélation des paramètres au sein d'une dynamique, il est important de connaître la corrélation des résidus entre les deux dynamiques. Une estimation est correcte de ce point de vue s'il n'existe pas de corrélation, c'est-à-dire que le coefficient doit être proche de zéro.

Le coefficient de corrélation, entre deux séries (X, Y) de même longueur N, est donné par le coefficient de Pearson. Il se base sur les observations, sa formule est :

$$\rho = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}$$

b) Estimation par le maximum de vraisemblance

L'estimation du modèle de Vasicek à deux facteurs par la méthode du maximum de vraisemblance va être réalisée aussi suivant les mêmes étapes que la méthode des moindres carrés, c'est-à-dire :

- Estimation des paramètres de la dynamique du taux long ;
- Estimation du taux long $\hat{l}_{t_{i+1}}$;
- Estimation des paramètres de la dynamique du taux court, à l'aide de $\hat{l}_{t_{i+1}}$.

(1) Estimation des paramètres du taux long

De la dynamique du taux long, on déduit que :

$$l_{t_{i+1}} \sim N \left[l_{t_i} \exp^{-a_l \delta} + \mu_l (1 - \exp^{-a_l \delta}), \frac{\sigma_l^2}{2a_l} (1 - \exp^{-2a_l \delta}) \right]$$

La fonction de vraisemblance est alors :

$$L(\theta_l) = \prod_{i=0}^{n-1} \frac{1}{\sigma_l \sqrt{\frac{1 - \exp^{-2a_l \delta}}{2a_l}} \sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{l_{t_{i+1}} - (\exp(-a_l \delta) l_{t_i} + \mu_l [1 - \exp(-a_l \delta)])}{\sigma_l \sqrt{\frac{1 - \exp^{-2a_l \delta}}{2a_l}}} \right)^2 \right)$$

La fonction de log vraisemblance devient alors :

$$\Lambda = -n \log \left(\sigma_l \sqrt{\frac{1 - \exp^{-2a_l \delta}}{2a_l}} \sqrt{2\pi} \right) - \frac{1}{2 \left(\sigma_l \sqrt{\frac{1 - \exp^{-2a_l \delta}}{2a_l}} \right)^2} \sum_{i=0}^{n-1} (l_{t_{i+1}} - \exp^{-2a_l \delta} l_{t_i} - \mu_l (1 - \exp^{-2a_l \delta}))^2$$

où $\Lambda = \sum_{i=0}^{n-1} \Lambda_{i+1}$ avec

$$\Lambda_{i+1} = -\log \left(\sigma_l \sqrt{\frac{1 - \exp^{-2a_l \delta}}{2a_l}} \sqrt{2\pi} \right) - \frac{1}{2 \left(\sigma_l \sqrt{\frac{1 - \exp^{-2a_l \delta}}{2a_l}} \sqrt{2\pi} \right)^2} (l_{t_{i+1}} - \exp^{-2a_l \delta} l_{t_i} - \mu_l (1 - \exp^{-2a_l \delta}))^2$$

Pour avoir le détail des calculs reportez vous à la partie sur le calibrage de l'inflation (cf. § L'inflation).

(2) Estimation du taux long $\hat{l}_{t_{i+1}}$

L'estimation des paramètres de la dynamique du taux court va s'appuyer sur les valeurs estimées du taux long. Pour cela on calcule le $\hat{l}_{t_{i+1}}$ de la façon suivante :

$$\hat{l}_{t_{i+1}} = \exp(-a_r \delta) l_{t_i} + \mu_{t_i} [1 - \exp(-a_r \delta)]$$

(3) Estimation des paramètres du taux court

En tenant compte de la valeur estimée du taux long, le taux court est :

$$r_{t_{i+1}} \sim N \left[r_{t_i} \exp^{-a_r \delta} + \hat{l}_{t_i} (1 - \exp^{-a_r \delta}), \frac{\sigma_r^2}{2a_r} (1 - \exp^{-2a_r \delta}) \right]$$

La fonction de vraisemblance :

$$L(\theta_r) = \prod_{i=0}^{n-1} \frac{1}{\frac{\sigma_r^2}{2a_r} (1 - \exp^{-2a_r \delta}) \sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{r_{t_{i+1}} - \exp^{-a_r \delta} r_{t_i} - \hat{l}_{t_i} (1 - \exp^{-a_r \delta})}{\frac{\sigma_r^2}{2a_r} (1 - \exp^{-2a_r \delta})} \right)^2 \right)$$

Et la fonction de log vraisemblance :

$$\Lambda = \sum_{i=0}^{n-1} \Lambda_{i+1}$$

avec

$$\Lambda_{i+1} = -\log \left(\sigma_r \sqrt{\frac{1 - \exp^{-2a_r \delta}}{2a_r}} \sqrt{2\pi} \right) - \frac{1}{2 \left(\sigma_r \sqrt{\frac{1 - \exp^{-2a_r \delta}}{2a_r}} \right)^2} \left(r_{t_{i+1}} - \exp^{-a_r \delta} r_{t_i} - \hat{l}_{t_i} (1 - \exp^{-a_r \delta}) \right)^2$$

Les résidus s'expriment sous la forme :

$$r_{i+1} = r_{t_{i+1}} - \exp^{-a_r \delta} r_{t_i} - \hat{l}_{t_i} (1 - \exp^{-a_r \delta})$$

Le gradient des résidus :

$$\begin{aligned} \frac{\partial r_{i+1}}{\partial a_r} &= \delta \exp^{-a_r \delta} (r_{t_i} - \hat{l}_{t_i}) \\ \frac{\partial r_{i+1}}{\partial \sigma_r} &= 0 \end{aligned}$$

Le gradient de la fonction log vraisemblance :

$$\begin{aligned} \frac{\partial \Lambda_{i+1}}{\partial a_r} &= -\frac{\delta \exp^{-2a_r \delta}}{1 - \exp^{-2a_r \delta}} + \frac{0.5}{a_r} - \frac{r_{i+1}^2}{\sigma_r^2 (1 - \exp^{-2a_r \delta})} - \frac{2a_r r_{i+1} \frac{\partial r_{i+1}}{\partial a_r}}{\sigma_r^2 (1 - \exp^{-2a_r \delta})} + \frac{2a_r \delta \exp^{-2a_r \delta} r_{i+1}^2}{\sigma_r^2 (1 - \exp^{-2a_r \delta})^2} \\ \frac{\partial \Lambda_{i+1}}{\partial \sigma_r} &= -\frac{1}{\sigma_r} + \frac{2a_r r_{i+1}^2}{\sigma_r^3 (1 - \exp^{-2a_r \delta})} \end{aligned}$$

La résolution des équations se fait en posant : $\alpha_r = \exp^{-a_r \delta}$ et $v_r^2 = \sigma_r^2 \frac{1 - \exp^{-2a_r \delta}}{2a_r}$, on

obtient alors :

$$r_{t_{i+1}} \sim N(r_{t_i} \alpha_r + \hat{l}_{t_i} (1 - \alpha_r), v_r^2)$$

La fonction de vraisemblance :

$$L(\theta_r) = \prod_{i=0}^{n-1} \frac{1}{v_r \sqrt{2\pi}} \exp \left(-\frac{(r_{i+1} - \alpha_r r_{t_i} - \hat{l}_{t_i}(1 - \alpha_r))^2}{2v_r^2} \right)$$

Et la fonction de log vraisemblance :

$$\Lambda = \sum_{i=0}^{n-1} \Lambda_{i+1} \text{ avec } \Lambda_{i+1} = -\log(v_r \sqrt{2\pi}) - \frac{1}{2v_r^2} r_{i+1}^2$$

Les résidus s'expriment sous la forme :

$$r_{i+1} = r_{t_{i+1}} - \alpha_r r_{t_i} - \hat{l}_{t_i}(1 - \alpha_r)$$

Le gradient des résidus :

$$\frac{\partial r_{i+1}}{\partial \alpha_r} = -r_{t_i} + \hat{l}_{t_i}$$

$$\frac{\partial r_{i+1}}{\partial v_r} = 0$$

Le gradient de log vraisemblance :

$$\frac{\partial \Lambda_{i+1}}{\partial \alpha_r} = -\frac{r_{i+1}}{v_r^2} \frac{\partial r_{i+1}}{\partial \alpha_r}$$

$$\frac{\partial \Lambda_{i+1}}{\partial v_r} = -\frac{1}{v_r} + \frac{r_{i+1}^2}{v_r^3}$$

La vraisemblance sera maximisée pour les paramètres vérifiant le système d'équation :

$$\begin{cases} \sum_{i=0}^{n-1} \left(\frac{\partial \Lambda_{i+1}}{\partial \alpha_r} \right) = -\sum_{i=0}^{n-1} \left(\frac{r_{i+1}}{v_r^2} \frac{\partial r_{i+1}}{\partial \alpha_r} \right) = 0 \\ \sum_{i=0}^{n-1} \left(\frac{\partial \Lambda_{i+1}}{\partial v_r} \right) = \sum_{i=0}^{n-1} \left(-\frac{1}{v_r} + \frac{r_{i+1}^2}{v_r^3} \right) = 0 \end{cases}$$

La résolution du système donne :

$$\bullet \quad \sum_{i=0}^{n-1} \left(\frac{\partial \Lambda_{i+1}}{\partial \alpha_r} \right) = -\sum_{i=0}^{n-1} \left(\frac{r_{i+1}}{v_r^2} \frac{\partial r_{i+1}}{\partial \alpha_r} \right) \Rightarrow -\sum_{i=0}^{n-1} (r_{t_{i+1}} - r_{t_i} \alpha_r - \hat{l}_{t_i}(1 - \alpha_r))(-r_{t_i} + \hat{l}_{t_i}) = 0$$

$$\hat{\alpha}_r = \frac{\sum_{i=0}^{n-1} r_{t_{i+1}} r_{t_i} + \sum_{i=0}^{n-1} r_{t_i} \hat{l}_{t_i} - \sum_{i=0}^{n-1} r_{t_{i+1}} \hat{l}_{t_i} - \sum_{i=0}^{n-1} \hat{l}_{t_i}^2}{\sum_{i=0}^{n-1} (r_{t_i} - \hat{l}_{t_i})^2}$$

$$\Rightarrow a_r = -\frac{\log(\hat{\alpha}_r)}{\delta}$$

$$\bullet \quad \hat{\sigma}_r = \sqrt{\frac{2\hat{\alpha}_r}{(n-2)(1 - \exp^{-2a_r \delta})} \frac{\sum_{i=0}^{n-1} r_{i+1}^2}{n}}$$

Matrice de variance – covariance des estimateurs

Déterminons la matrice Hessienne pour en déduire la matrice de variance, covariance.

Le gradient des résidus :

$$\frac{\partial r_{i+1}}{\partial a_r} = \delta \exp^{-a_r \delta} (r_{t_i} - \hat{l}_{t_i})$$

$$\frac{\partial r_{i+1}}{\partial \sigma_r} = 0$$

Pour la log vraisemblance :

$$\begin{aligned} \frac{\partial \Lambda_{i+1}}{\partial a_r} &= -\frac{\delta \exp^{-2a_r \delta}}{1 - \exp^{-2a_r \delta}} + \frac{0.5}{a_r} - \frac{r_{i+1}^2}{\sigma_r^2 (1 - \exp^{-2a_r \delta})} - \frac{2a_r r_{i+1} \frac{\partial r_{i+1}}{\partial a_r}}{\sigma_r^2 (1 - \exp^{-2a_r \delta})} + \frac{2a_r \delta \exp^{-2a_r \delta} r_{i+1}^2}{\sigma_r^2 (1 - \exp^{-2a_r \delta})^2} \\ \frac{\partial \Lambda_{i+1}}{\partial \sigma_r} &= -\frac{1}{\sigma} + \frac{2a_r r_{i+1}^2}{\sigma^3 (1 - \exp^{-2a_r \delta})} \end{aligned}$$

Les dérivées secondes des résidus :

$$\frac{\partial^2 r_{i+1}}{\partial a_r^2} = -\delta^2 \exp^{-a_r \delta} (r_{t_i} - \hat{l}_{t_i}) ;$$

$$\frac{\partial^2 r_{i+1}}{\partial a_r \partial \sigma_r} = \frac{\partial^2 r_{i+1}}{\partial \sigma_r^2} = 0.$$

Les dérivées secondes de la log vraisemblance :

$$\begin{aligned} \frac{\partial^2 \Lambda_{i+1}}{\partial a_r^2} &= \frac{2\delta^2 \exp^{-2a_r \delta}}{1 - \exp^{-2a_r \delta}} + \frac{2\delta^2 \exp^{-4a_r \delta}}{(1 - \exp^{-2a_r \delta})^2} - \frac{0.5}{a_r^2} - \frac{2r_{i+1} \frac{\partial r_{i+1}}{\partial a_r}}{\sigma_r^2 (1 - \exp^{-2a_r \delta})} + \frac{2\delta \exp^{-2a_r \delta} r_{i+1}^2}{\sigma_r^2 (1 - \exp^{-2a_r \delta})^2} \\ &+ \frac{2\delta \exp^{-2a_r \delta} r_{i+1}^2}{\sigma_r^2 (1 - \exp^{-2a_r \delta})^2} - \frac{4a_r \delta^2 \exp^{-2a_r \delta} r_{i+1}^2}{\sigma_r^2 (1 - \exp^{-2a_r \delta})^2} + \frac{4a_r \delta \exp^{-2a_r \delta} r_{i+1} \frac{\partial r_{i+1}}{\partial a_r}}{\sigma_r^2 (1 - \exp^{-2a_r \delta})^2} - \frac{8a_r \delta^2 \exp^{-4a_r \delta} r_{i+1}^2}{\sigma_r^2 (1 - \exp^{-2a_r \delta})^3} \\ &- \frac{2r_{i+1} \frac{\partial r_{i+1}}{\partial a_r}}{\sigma_r^2 (1 - \exp^{-2a_r \delta})} - \frac{2a_r r_{i+1} \frac{\partial^2 r_{i+1}}{\partial a_r^2}}{\sigma_r^2 (1 - \exp^{-2a_r \delta})} - \frac{2a_r \left(\frac{\partial r_{i+1}}{\partial a_r} \right)^2}{\sigma_r^2 (1 - \exp^{-2a_r \delta})} + \frac{4a_r \delta \exp^{-2a_r \delta} r_{i+1} \frac{\partial r_{i+1}}{\partial a_r}}{\sigma_r^2 (1 - \exp^{-2a_r \delta})^2} \end{aligned}$$

$$\frac{\partial^2 \Lambda_{i+1}}{\partial a_r \partial \sigma_r} = \frac{2r_{i+1}^2}{\sigma_r^3 (1 - \exp^{-2a_r \delta})} + \frac{4a_r r_{i+1} \frac{\partial r_{i+1}}{\partial a_r}}{\sigma_r^3 (1 - \exp^{-2a_r \delta})} - \frac{4a_r \delta \exp^{-2a_r \delta} r_{i+1}^2}{\sigma_r^3 (1 - \exp^{-2a_r \delta})^2}$$

$$\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma_r \partial a_r} = \frac{2r_{i+1}^2}{\sigma_r^3 (1 - \exp^{-2a_r \delta})} + \frac{4a_r r_{i+1} \frac{\partial r_{i+1}}{\partial a_r}}{\sigma_r^3 (1 - \exp^{-2a_r \delta})} - \frac{4a_r \delta \exp^{-2a_r \delta} r_{i+1}^2}{\sigma_r^3 (1 - \exp^{-2a_r \delta})^2}$$

$$\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma_r^2} = \frac{1}{\sigma_r^2} - \frac{6a_r r_{i+1}^2}{\sigma_r^4 (1 - \exp^{-2a_r \delta})}$$

Il ne faut pas oublier que $\Lambda = \sum_{i=0}^{n-1} \Lambda_{i+1}$, la matrice Hessienne est :

$$H_{\Lambda}(\theta_r) = \begin{pmatrix} \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial a_r^2} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial a_r \partial \sigma_r} \right) \\ \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma_r \partial a_r} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma_r^2} \right) \end{pmatrix}$$

(4) Corrélation entre les résidus des dynamiques

Le coefficient de corrélation, entre deux séries (X, Y) de même longueur N, est donné par le coefficient de Pearson :

$$\rho = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}$$

3. Résultat du calibrage

a) Méthode d'estimation retenue

De la même façon que pour le modèle de Vasicek à un facteur, les méthodes d'estimations donnent des valeurs de paramètres et des vecteurs de résidus identiques. La méthode du maximum de vraisemblance est retenue car elle donne plus d'information.

b) 1^{er} test : 31 données (1978-2009)

Valeur des paramètres estimés

Pour le taux long :

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Force de Rappel	0,2483	0,1128	0,0361	La force de rappel est significative
Moyenne	0,0114	0,005	0,0301	La moyenne est significative
Volatilité	0,0094	0,0013	<0.0001	La volatilité est significative

Pour le taux court :

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Force de Rappel	0,4162	0,2217	0,071	La force de rappel est presque significative
Volatilité	0,0139	0,0022	<0.00001	La volatilité est significative

Qualité d'ajustement du modèle

L'analyse de la variance

Pour le taux long :

Source	ddl	SCE	CM	F	Prob>F
Modèle	1	0,042418	0,042418	565,984	0
Erreur	28	0,002098	7,5E-05		
Total	29	0,044516			

Pour le taux court :

Source	ddl	SCE	CM	F	Prob>F
Modèle	1	0,003676	0,003676	26,237	0,00002
Erreur	28	0,003923	0,00014		
Total	29	0,0076			

Les coefficients de détermination ajustés :

Le $R^2_{ajusté}$ pour le taux long est de 0,9512 et pour le taux court, il est de 0,4653. Le pouvoir explicatif du modèle à taux court est satisfaisant contrairement à celui du taux long.

La matrice de corrélation entre les estimateurs

Pour le taux long :

	Force de Rappel	Moyenne	Volatilité
Force de Rappel	1	0,9388	0,372
Moyenne	0,9388	1	0,3492
Volatilité	0,372	0,3492	1

Pour le taux court :

	Force de Rappel	Volatilité
Force de Rappel	1	0,5954
Volatilité	0,5954	1

Test d'hypothèse du modèle

Pour le taux long :

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Moyenne = 0	0	1	La moyenne des résidus est considérée nulle
Variance = 1	30	0,414	La variance des résidus est considérée valoir 1
Shapiro-Wilk	0,9808	0,8471	Hypothèse de normalité des résidus vérifiée
Breush-Godfrey	13,4593	0,0002	Présence d'autocorrélation d'ordre 1
Arch(1) LM	0,1622	0,6871	Absence d'hétéroscédacité d'ordre 1

Pour le taux court :

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Moyenne = 0	-0,0031	0,7851	La moyenne des résidus est considérée nulle
Variance = 1	29,9293	0,4175	La variance des résidus est considérée valoir 1
Shapiro-Wilk	0,9799	0,8218	Hypothèse de normalité des résidus vérifiée
Breush-Godfrey	6,022	0,0141	Présence d'autocorrélation d'ordre 1
Arch(1) LM	0,3728	0,5415	Absence d'hétéroscédacité d'ordre 1

Corrélation entre les résidus :

La corrélation entre les résidus est de 0,1109 donc proche de zéro, ce qui était souhaitable pour obtenir un modèle de qualité.

c) 2^{ème} test : toute la série de données (1951-2009)

Le premier calibrage donnant des résultats pas totalement satisfaisant concernant la significativité du paramètre de la force de rappel du taux court, un second test est réalisé en prenant toute la série des données. Les résultats obtenus sont présentés ci-après.

Valeur des paramètres estimés

Pour le taux long :

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Force de Rappel	0,6938	0,1937	0,0007	La force de rappel est significative
Moyenne	0,0225	0,007	0,002	La moyenne est significative
Volatilité	0,0288	0,0034	<0.0001	La volatilité est significative

Mise en œuvre du calibrage des classes d'actifs

Pour le taux court :

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Force de Rappel	0,9002	0,3252	0,0076	La force de rappel est significative
Volatilité	0,0339	0,005	<0.00001	La volatilité est significative

Qualité d'ajustement du modèle

L'analyse de la variance

Pour le taux long :

Source	ddl	SCE	CM	F	Prob>F
Modèle	1	0,028668	0,028668	61,918	0
Erreur	56	0,025928	0,000463		
Total	57	0,054596			

Pour le taux court :

Source	ddl	SCE	CM	F	Prob>F
Modèle	1	0,007959	0,007959	14,911	0,00029
Erreur	56	0,02989	0,000534		
Total	57	0,037849			

Les coefficients de détermination ajustés :

Le $R^2_{ajusté}$ pour le taux long est de 0,5166 et pour le taux court, il est de 0,1962. Le pouvoir explicatif des modèles n'est pas satisfaisant.

La matrice de corrélation entre les estimateurs

Pour le taux long :

	Force de Rappel	Moyenne	Volatilité
Force de Rappel	1	0,833	0,629
Moyenne	0,833	1	0,5239
Volatilité	0,629	0,5239	1

Pour le taux court :

	Force de Rappel	Volatilité
Force de Rappel	1	0,7814
Volatilité	0,7814	1

Test d'hypothèse du modèle

Pour le taux long :

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Moyenne = 0	0	1	La moyenne des résidus est considérée nulle
Variance = 1	58	0,4382	La variance des résidus est considérée valoir 1
Shapiro-Wilk	0,7776	0	Hypothèse de normalité des résidus non vérifiée
Breush-Godfrey	1,5708	0,2101	Absence d'autocorrélation d'ordre 1
Arch(1) LM	0,0378	0,8459	Absence d'hétéroscédacité d'ordre 1

Pour le taux court :

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Moyenne = 0	-0,007	0,7605	La moyenne des résidus est considérée nulle

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Variance = 1	56,2521	0,5031	La variance des résidus est considérée valoir 1
Shapiro-Wilk	0,8493	0	Hypothèse de normalité des résidus non vérifiée
Breush-Godfrey	6,6508	0,0099	Présence d'autocorrélation d'ordre 1
Arch(1) LM	0,0123	0,9118	Absence d'hétéroscédacité d'ordre 1

La corrélation entre les résidus est de -0,0062, donc proche de 0.

d) Conclusion

La significativité des paramètres est constatée lorsqu'on utilise toute la série des données pour calibrer le modèle, sinon la force de rappel du taux court est à peine significative. En revanche, la qualité d'ajustement des modèles s'est fortement dégradée dans le second test. Dans les deux cas, la corrélation entre les paramètres estimés est forte. La corrélation entre les résidus est quant à elle très faible, et même négative dans le deuxième test.

En étudiant de plus près la valeur des paramètres estimés :

- La force de rappel est plus importante sur le taux court
- La volatilité est plus importante sur le court terme que le long terme

Ceci correspond à ce l'on attendait.

Enfin, concernant les tests d'hypothèses, tous ne sont pas vérifiés, quand on calibre sur l'ensemble de la série, il semblerait que cela se dégrade car trois tests sont non vérifiés contre deux pour l'estimation sur 30 données.

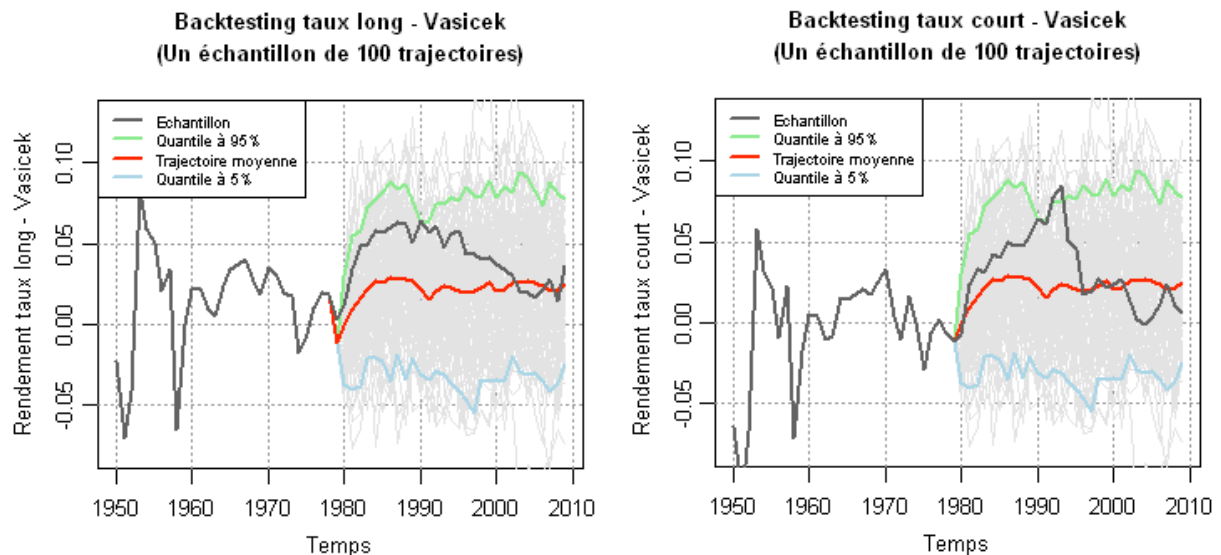
Remarque : de même que pour Vasicek à un facteur, ce ne sont pas les résidus bruts qui sont étudiés pour les tests d'hypothèses.

Le calibrage qu'il serait préférable de retenir ne s'impose pas de façon évidente. Etudions les critères de comparaison des modèles :

Vasicek 2F	Taux court		Taux long	
	Test 1	Test 2	Test 1	Test 2
N	31	59	31	59
K	2	2	3	3
Log vraisemblance	91,5264	128,9755	100,9480	141,3747
Critère AICc	0,5055	-0,0067	2,8807	2,1356
Critère BIC	-0,9404	-0,6793	-1,0255	-0,7590

Comme pour l'inflation, il est préférable de retenir le calibrage effectué sur l'ensemble de la série, le AICc étant plus faible dans ce cas.

Le backtesting sur le calibrage du modèle sur la série complète montre qu'on ne surestime pas les taux, ce qui est souvent le cas de l'estimation par maximum de vraisemblance :



C. Les actions

Le GSE d'Ahlgrim propose le modèle d'Hardy à deux régimes pour modéliser les actions hors dividendes, les données de Friggit dont nous disposons sont les valeurs des actions avec dividendes réinvestis, ce qui ne permet pas de calibrer correctement le modèle.

De plus en imaginant qu'on puisse estimer les paramètres du modèle, comme savoir si le modèle est de bonne qualité ?

Nous choisissons d'utiliser le modèle très répandu de Black & Scholes (noté B&S) en substitution au modèle de Hardy, cela correspond en fait à un modèle de Hardy avec un seul régime. Ce modèle est simple d'implémentation, ce qui en fait sa force, mais aussi sa faiblesse.

Les principaux inconvénients souvent mis en exergue pour le modèle B&S :

- Les trajectoires sont continues, ce qui exclut les sauts de valeur, les chocs possibles sur la valeur des actions sont sous-estimés;
- La volatilité définie est constante, ce qui n'est pas le cas empiriquement ;
- Les rendements des actions suivent une loi log-normale¹⁵, ce qui n'est pas toujours vérifié empiriquement ;
- Les queues de distribution de la loi (log-)normale ne sont pas suffisamment épaisses par rapport à la réalité. Ce qui veut dire que certaines valeurs extrêmes donc certains risques sont moins pris en compte.

Au vu des critiques du modèle B&S, nous testerons le calibrage des séries de données actions en utilisant les modèles de B&S et de Merton qui permet de prendre en compte les sauts de valeur.

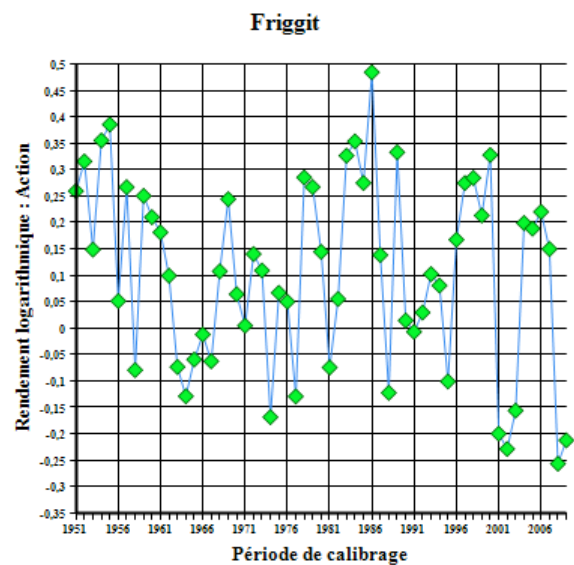
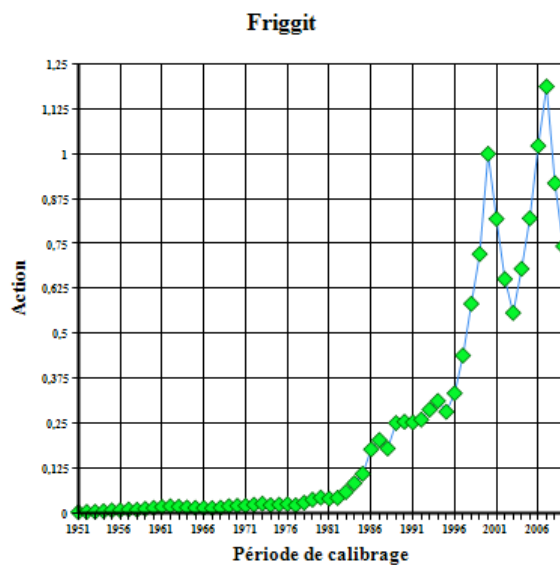
¹⁵ Loi log-normale : une variable aléatoire X suit une loi log-normale de paramètres μ et σ si la variable $Y=\ln(X)$ suit une loi normale de paramètres μ et σ .

1. Les données

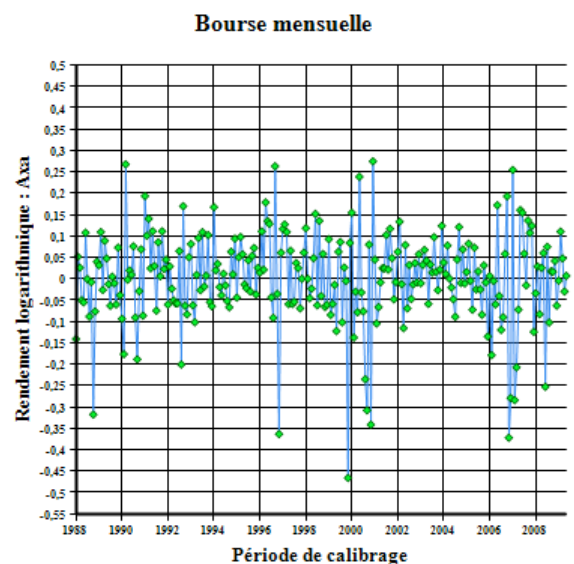
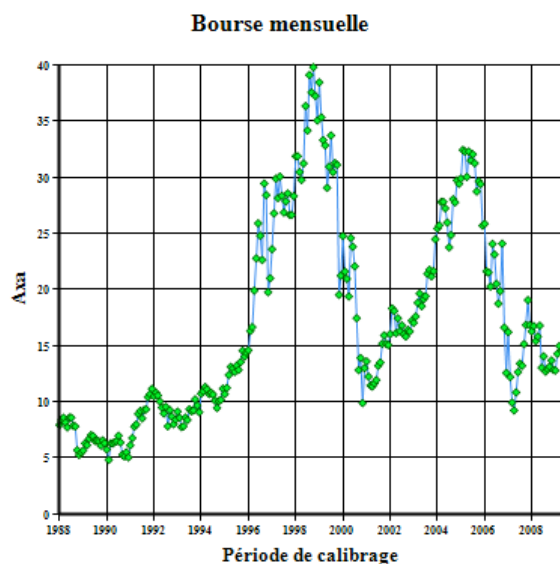
Deux séries de données sont retenues pour réaliser un calibrage. La première série est celle des données de Friggit, avec une fréquence annuelle. La seconde série correspond à l'action AXA, les données sont mensuelles et issues de la bourse.

L'hypothèse retenue est de modéliser les rendements actions avec dividendes réinvestis. Pour calibrer le modèle, la première chose à faire est de rendre les données exploitables en calculant :

$$\text{« Rendement Action »}_i = \text{Log Népérien} (\text{Action}_i / \text{Action}_{i-1})$$



A priori, graphiquement, il n'y aurait pas de saut dans les données de Friggit mais plutôt des pics de valeurs.



L'action AXA semble avoir des sauts dans sa courbe, reste à trouver le modèle qui sera le mieux adapté à ce cas.

2. Modèle de Black & Scholes

La dynamique

Le prix de l'action, noté $S(t)$, suit un mouvement brownien géométrique, défini comme suit :

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t$$

Avec μ la tendance de l'action

σ la volatilité de l'action

B_t mouvement brownien standard

La solution de l'équation nous donne le prix des actions :

$$S(t) = S(0) \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right)$$

Avec $S(0)$ le prix initial de l'action.

La méthode d'estimation retenue est celle du maximum de vraisemblance.

Maximum de vraisemblance

Les rendements de l'action sont définis par : $r_{t_{i+1}} = \log\left(\frac{S_{t_{i+1}}}{S_{t_i}}\right)$

$$r_{t_{i+1}} = \log\left(\frac{S_{t_i} \exp\left((\mu - \sigma^2/2)(t_{i+1} - t_i) + \sigma(B_{t_{i+1}} - B_{t_i})\right)}{S_{t_i}}\right) = (\mu - \frac{\sigma^2}{2})(t_{i+1} - t_i) + \sigma(B_{t_{i+1}} - B_{t_i})$$

D'où $r_{t_{i+1}} \sim N((\mu - \frac{\sigma^2}{2})\delta, \sigma^2\delta)$ avec δ le pas de temps.

La vraisemblance s'écrit : $L(\theta) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2\delta}} \exp\left(-\frac{(r_{t_i} - (\mu - \frac{\sigma^2}{2})\delta)^2}{2\sigma^2\delta}\right)$

Le principe de calibrage pour le modèle de Black et Scholes est exactement le même que celui de Vasicek, le détail des calculs ne sera pas déroulé.

Le rendement des actions est ensuite donné par :

$$r(t) = \frac{S(t) - S(t-1)}{S(t-1)}$$

3. Modèle de Merton

Les calculs permettant le calibrage du modèle de Merton ont été développés dans la partie précédente, vous pouvez vous y reporter pour plus d'information (cf. §Particularité d'un modèle intégrant des sauts : Merton).

4. Résultat du calibrage – série de Friggit

a) 1^{er} test : 31 données (1978-2009) – B&S

Le premier test va utiliser le modèle de B&S sur une trentaine d'année.

Valeur des paramètres estimés

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Tendance	0,1251	0,036	0,0017	La tendance est significative
Volatilité	0,1985	0,0258	<0.00001	La volatilité est significative

Qualité d'ajustement du modèle

La matrice de corrélation entre les estimateurs :

	Tendance	Volatilité
Tendance	1	0,1424
Volatilité	0,1424	1

Test d'hypothèse du modèle

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Moyenne = 0	0	1	La moyenne des résidus est considérée nulle
Variance = 1	30	0,4657	La variance des résidus est considérée valoir 1
Shapiro-Wilk	0,9575	0,2509	Hypothèse de normalité des résidus vérifiée
Breush-Godfrey	3,9895	0,0458	Présence d'autocorrélation d'ordre 1
Arch(1) LM	1,3718	0,2415	Absence d'hétéroscédacité d'ordre 1

b) 2^{ème} test : toute la série de données (1951-2009) – B&S

Valeur des paramètres estimés

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Tendance	0,1201	0,0234	<0.00001	La tendance est significative
Volatilité	0,1785	0,0166	<0.00001	La volatilité est significative

Qualité d'ajustement du modèle

La matrice de corrélation entre les estimateurs :

	Tendance	Volatilité
Tendance	1	0,1269
Volatilité	0,1269	1

Test d'hypothèse du modèle

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Moyenne = 0	0	1	La moyenne des résidus est considérée nulle
Variance = 1	58	0,4753	La variance des résidus est considérée valoir 1
Shapiro-Wilk	0,9759	0,29	Hypothèse de normalité des résidus vérifiée
Breush-Godfrey	6,5001	0,0108	Présence d'autocorrélation d'ordre 1
Arch(1) LM	1,9779	0,1596	Absence d'hétéroscédacité d'ordre 1

c) 3^{ème} test : toute la série de données (1988 -2009) – Merton

Valeur des paramètres estimés

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Tendance	0,1218	0,0254	0	La tendance est significative
Volatilité	0,188	0,02	<0.00001	La volatilité est significative
Indicateur de Saut	0,056	Non	Non	L'indicateur de Saut n'est pas significatif
Volatilité du Saut	0,2676	Non	Non	La volatilité du Saut n'est pas significative
		Numérique	Numérique	
		Numérique	Numérique	

Tests d'ajustement et d'hypothèse

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
iid	35	1,70843	Le caractère iid est vérifié.
Khi2	0,9156	.	L'adéquation au modèle est vérifiée.

d) Conclusion

La modélisation par le modèle de Merton ne semble pas adaptée à la série de données de Friggit. En considérant toute la série de données, l'indicateur de saut et sa volatilité ne sont pas significatifs, ceci signifie qu'il n'y a pas de saut dans la série, c'est ce qui était pressenti en observant les données.

Ceci est confirmé par le test du rapport de vraisemblance entre les modèles. La vraisemblance pour le modèle de B&S est 18,4344, pour Merton, elle est de 17,6335, d'où le rapport de vraisemblance à $-2 \cdot (18,4344 - 17,6335) = -1,6018$. La p_value est de 1, l'apport du modèle de Merton n'est pas significatif.

Nous retenons alors la modélisation par B&S. Les paramètres estimés sont significatifs dans les deux cas de tests menés en utilisant le modèle B&S.

Il est intéressant de remarquer que les valeurs sont globalement doublées entre les deux tests, même pour la corrélation des paramètres.

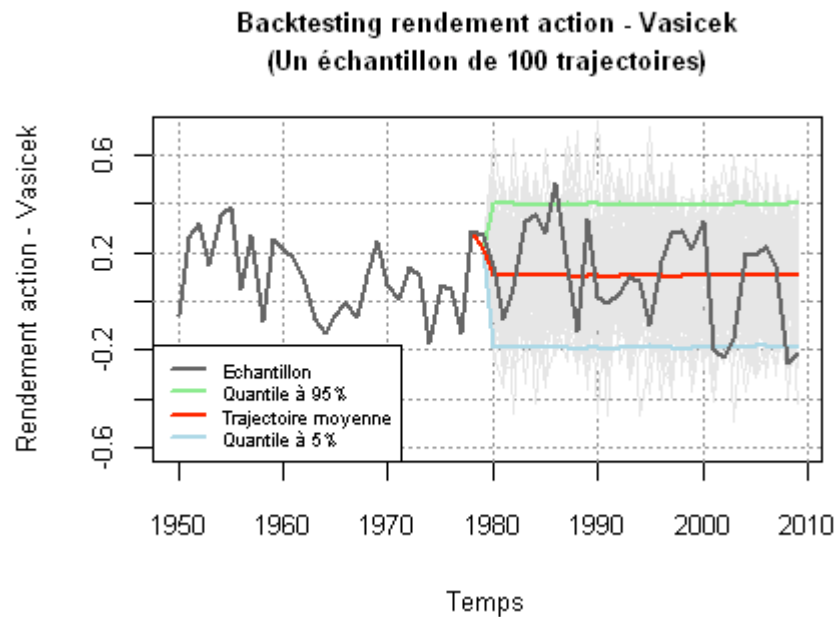
Concernant les tests d'hypothèses, il existe dans les deux cas de l'autocorrélation.

Le calibrage suivant l'une ou l'autre des séries de données offre une qualité similaire du modèle. Calculons les critères AICc et BIC :

	Test 1	Test 2
N	31	59
K	2	2
Log vraisemblance	6,6433	18,4344
Critère AICc	2,7838	1,6830
Critère BIC	1,3380	1,0104

En utilisant le critère AICc, c'est une fois de plus le calibrage sur la plus grande période qui semble le meilleur.

Backtesting sur le calibrage long terme :



Les résultats de l'estimation par backtesting semblent corrects.

5. Résultat du calibrage – série d'AXA

a) 1^{er} test : toute la série de données (1951-2009) – B&S

Valeur des paramètres estimés

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Tendance	0,0006	0,0006	0,2612	La tendance n'est pas significative
Volatilité	0,0307	0,0014	<0.00001	La volatilité est significative

La matrice de corrélation entre les estimateurs

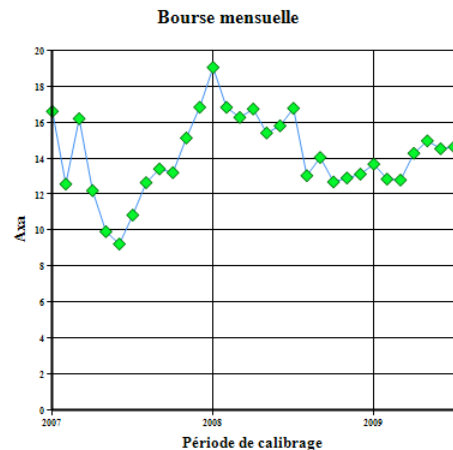
	Tendance	Volatilité
Tendance	1	0,0752
Volatilité	0,0752	1

Test d'hypothèse du modèle

Test $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Moyenne = 0	0,0017	0,9874	La moyenne des résidus est considérée nulle
Variance = 1	3072	0	La variance des résidus n'est pas considérée valoir 1
Shapiro-Wilk	0,9442	0	Hypothèse de normalité des résidus non vérifiée
Breush-Godfrey	1,2313	0,2671	Absence d'autocorrélation d'ordre 1
Arch(1) LM	9,2612	0,0023	Présence d'hétéroscédacité d'ordre 1

b) 2^{ième} test : 31 données (2007-2009) – Merton

Ce calibrage correspond à la partie de la courbe suivante :



Il semble qu'il y ait plusieurs sauts dans ce petit historique de données. En évaluant les paramètres, on a :

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Tendance	0,0901	0,1994	0,655	La tendance n'est pas significative
Volatilité	0,1321	Non	Non	La volatilité n'est pas significative
Indicateur de Saut	21,6375	9,9227	0,0378	L'indicateur de Saut est significatif
Volatilité du Saut	0,1049	0,039	0,012	La volatilité du Saut est significative

Tests d'ajustement et d'hypothèse

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
iid	21	0,55143	Le caractère iid est vérifié.
Khi2	2,3819	0,9925	L'adéquation au modèle est vérifiée.

c) 3^{ème} test : toute la série de données (1988 -2009) – Merton

Valeur des paramètres estimés

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Tendance	0,1377	0,0659	0,0376	La tendance est significative
Volatilité	0,2424	0,0195	<0.00001	La volatilité est significative
Indicateur de Saut	2,8772	1,1751	0,015	L'indicateur de Saut est significatif
Volatilité du Saut	0,1652	0,0327	<0.00001	La volatilité du Saut est significative

Matrice de corrélation des estimateurs

	Tendance	Volatilité	Indicateur de Saut	Volatilité du Saut
Tendance	1	0,1502	-0,1844	0,1811
Volatilité	0,1502	1	-0,5904	0,3993
Indicateur de Saut	-0,1844	-0,5904	1	-0,7641
Volatilité du Saut	0,1811	0,3993	-0,7641	1

Tests d'ajustement et d'hypothèse

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
iid	173	0,69159	Le caractère iid est vérifié.
Khi2	3,0831	0,9794	L'adéquation au modèle est vérifiée.

d) Conclusion

Dans le premier test, le modèle de Black et Scholes ne répond pas correctement à la modélisation de la série de données AXA, même en prenant en compte l'ensemble des données, la tendance n'est pas significative. De plus, les résidus ne suivent pas une loi normale, B&S n'est alors pas applicable, c'est une critique fréquente du modèle.

Les tests s'appuyant sur le modèle de Merton sont plus concluants.

Le second test sur 31 données révèle que seuls les paramètres de saut sont significatifs, le modèle contient bien des sauts mais avec si peu de données, il n'est pas possible d'avoir un bon calibrage.

Enfin, en considérant l'ensemble des données Axa, sur le dernier test, le modèle de Merton est bon, les paramètres sont significatifs et les tests validés.

L'estimateur de l'indicateur de saut nous permet de dire qu'il y a en moyenne 2,8 sauts par an dans la série de données.

Remarque : dans la matrice de corrélation, on constate que l'indicateur de saut évolue en sens inverse de la volatilité, en effet au moment d'un saut, la volatilité du modèle est principalement portée par les paramètres du saut, ce qui diminue mécaniquement la volatilité sur le reste du modèle, cela semble logique.

Le test du rapport de vraisemblance permet d'être conforté dans le choix du modèle. La vraisemblance pour le modèle de B&S est -877,431, pour Merton, elle est de 230,4767, d'où le rapport de vraisemblance à $-2*(-877,431-230,4767)=2215,8153$. La p_value est de 0,00001, l'apport du modèle de Merton est significatif.

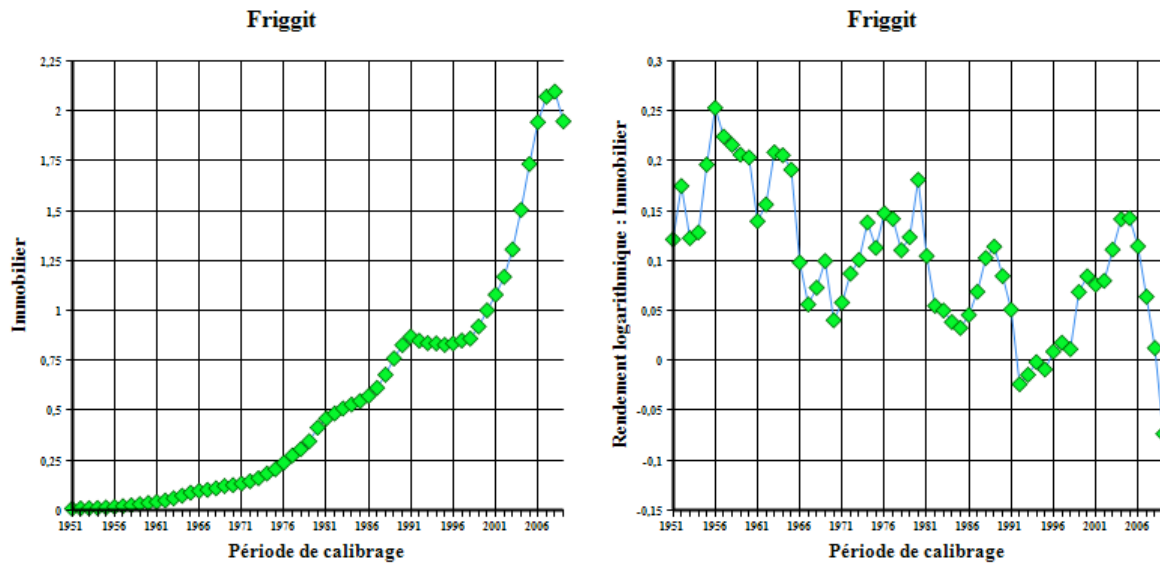
Le calibrage à retenir pour la série de Friggitt est B&S sur 31 données, en revanche pour l'action AXA, il est préférable d'avoir le modèle de Merton calibré sur l'ensemble de la série.

D. L'immobilier

1. Les données

Les données de l'immobilier de Friggitt sont sous leurs formes brutes, pour les rendre exploitables nous calculons les rendements logarithmiques de la façon suivante :

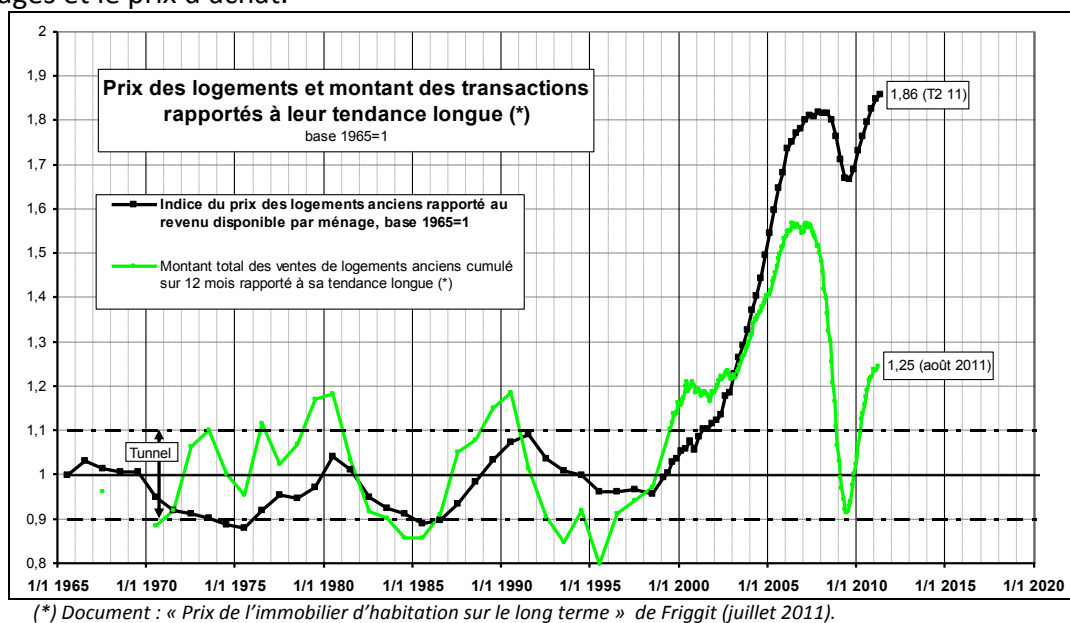
$$\text{« Rendement Immobilier »}_i = \text{Log Népérien} (\text{Immobilier}_i / \text{Immobilier}_{i-1})$$



Le premier graphique montre que les prix de l'immobilier ont fortement augmenté à partir du milieu des années 90. L'étude plus détaillée des données et des travaux de Friggit, nous permet d'identifier plusieurs phénomènes :

- Une forte augmentation des prix de l'immobilier depuis le début des années 2000 ;
- A partir de 2002, le rapport entre l'indice de prix des logements et le revenu des ménages devient très important au point de sortir du tunnel de Friggit (considéré comme l'intervalle de valeur « normale » de l'immobilier) ;
- Depuis 2005/2006, on peut considérer qu'étant largement sortis du tunnel de Friggit, nous sommes en présence d'une bulle immobilière.

De grands économistes et spécialistes de l'immobilier tel que Friggit prévoient un éclatement de la bulle immobilière dans un avenir très proche, en raison d'une déconnexion entre d'une part le prix des loyers et le prix d'achat, et d'autre part entre le revenu des ménages et le prix d'achat.



Cet avis n'est cependant pas unanime, certains pensant que la progression des valeurs de l'immobilier va se poursuivre comme l'économiste M. Mouillart.

Dans ce contexte, comment calibrer correctement le modèle et évaluer les titres immobiliers en valeur de marché ?

2. Modèle de Vasicek

Le modèle de Vasicek a déjà été utilisé pour calibrer l'inflation, les calculs ne sont pas redéveloppés ici, vous pouvez vous reporter à la partie concernée pour plus de détail (cf. §L'inflation).

Nous présentons directement les résultats du calibrage.

3. Modèle de Black & Scholes

Le modèle de Black & Scholes a déjà été utilisé pour calibrer les actions, vous pouvez vous reporter à la partie concernée pour plus de détail (cf. §Les actions).

4. Résultat du calibrage

a) 1^{er} test : 31 données (1978-2009) - Vasicek

Valeur des paramètres estimés

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Force de Rappel	0,1791	0,1521	0,2489	La force de Rappel n'est pas significative
Moyenne	0,0043	0,0116	0,7117	La moyenne n'est pas significative
Volatilité	0,0386	0,0057	<0.00001	La volatilité est significative

Qualité d'ajustement du modèle

L'analyse de la variance :

Source	ddl	SCE	CM	F	Prob>F
Modèle	1	0,054262	0,054262	40,351	0
Erreur	28	0,037653	0,001345		
Total	29	0,091916			

Le coefficient de détermination ajusté :

Le $R^2_{ajusté}$ est de 0,5757, le pouvoir explicatif du modèle n'est pas satisfaisant.

La matrice de corrélation entre les estimateurs :

	Force de Rappel	Moyenne	Volatilité
Force de Rappel	1	0,7949	0,4846
Moyenne	0,7949	1	0,3852
Volatilité	0,4846	0,3852	1

Test d'hypothèse du modèle

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Moyenne = 0	0	1	La moyenne des résidus est considérée nulle
Variance = 1	30	0,414	La variance des résidus est considérée valoir 1
Shapiro-Wilk	0,9814	0,8605	Hypothèse de normalité des résidus vérifiée
Breusch-Godfrey	5,177	0,0229	Présence d'autocorrélation d'ordre 1
Arch(1) LM	1,2051	0,2723	Absence d'hétéroscédacité d'ordre 1

b) 2^{ème} test : toute la série de données (1951-2009) - Vasicek

Valeur des paramètres estimés

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Force de Rappel	0,1091	0,0819	0,1886	La force de Rappel n'est pas significative
Moyenne	0,0076	0,0097	0,4365	La moyenne n'est pas significative
Volatilité	0,0392	0,004	<0.00001	La volatilité est significative

Qualité d'ajustement du modèle

L'analyse de la variance :

Source	ddl	SCE	CM	F	Prob>F
Modèle	1	0,205884	0,205884	143,772	0
Erreur	56	0,080193	0,001432		
Total	57	0,286077			

Le coefficient de détermination ajusté :

Le $R^2_{ajusté}$ est de 0,7147, le pouvoir explicatif du modèle n'est pas satisfaisant.

La matrice de corrélation entre les estimateurs :

	Force de Rappel	Moyenne	Volatilité
Force de Rappel	1	0,848	0,3914
Moyenne	0,848	1	0,3319
Volatilité	0,3914	0,3319	1

Test d'hypothèse du modèle

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Moyenne = 0	0	1	La moyenne des résidus est considérée nulle
Variance = 1	58	0,4382	La variance des résidus est considérée valoir 1
Shapiro-Wilk	0,9775	0,3551	Hypothèse de normalité des résidus vérifiée
Breush-Godfrey	3,8588	0,0495	Présence d'autocorrélation d'ordre 1
Arch(1) LM	1,0246	0,3114	Absence d'hétéroscédacité d'ordre 1

c) 3^{ème} test : 31 données (1978-2009) – B&S

Valeur des paramètres estimés

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Tendance	0,0594	0,0103	<0.00001	La tendance est significative
Volatilité	0,0563	0,0075	<0.00001	La volatilité est significative

Qualité d'ajustement du modèle

La matrice de corrélation entre les estimateurs :

	Tendance	Volatilité
Tendance	1	0,0408
Volatilité	0,0408	1

Test d'hypothèse du modèle

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Moyenne = 0	0	1	La moyenne des résidus est considérée nulle

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Variance = 1	29	0,4651	La variance des résidus est considérée valoir 1
Shapiro-Wilk	0,9929	0,999	Hypothèse de normalité des résidus vérifiée
Breush-Godfrey	14,2359	0,0002	Présence d'autocorrélation d'ordre 1
Arch(1) LM	1,0911	0,2962	Absence d'hétéroscédacité d'ordre 1

d) 4^{ème} test : toute la série de données (1951-2009) – B&S

Valeur des paramètres estimés

Estimateur	Valeur	Ecart type	Probabilité	Interprétation
Tendance	0,1019	0,0092	<0.00001	La Tendance est significative
Volatilité	0,0703	0,0066	<0.00001	La Volatilité est significative

Qualité d'ajustement du modèle

La matrice de corrélation entre les estimateurs :

	Tendance	Volatilité
Tendance	1	0,0503
Volatilité	0,0503	1

Test d'hypothèse du modèle

Test pour $\alpha = 5\%$	Statistique	Probabilité	Interprétation
Moyenne = 0	0	1	La moyenne des résidus est considérée nulle
Variance = 1	58	0,4753	La variance des résidus est considérée valoir 1
Shapiro-Wilk	0,9895	0,8911	Hypothèse de normalité des résidus vérifiée
Breush-Godfrey	42,3649	0	Présence d'autocorrélation d'ordre 1
Arch(1) LM	20,798	0	Présence d'hétéroscédacité d'ordre 1

e) Conclusion

Modèle de Vasicek

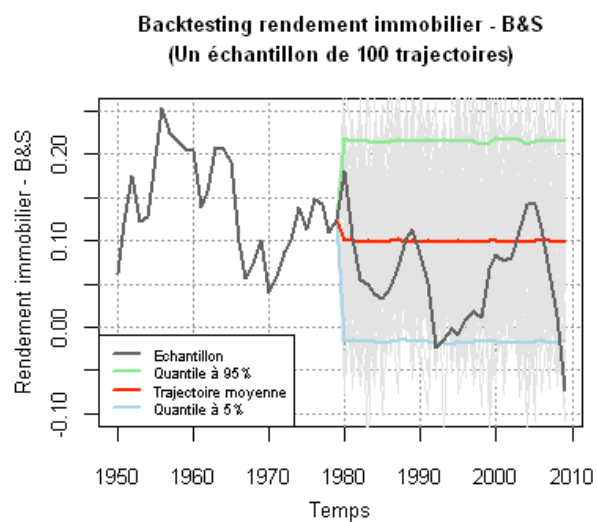
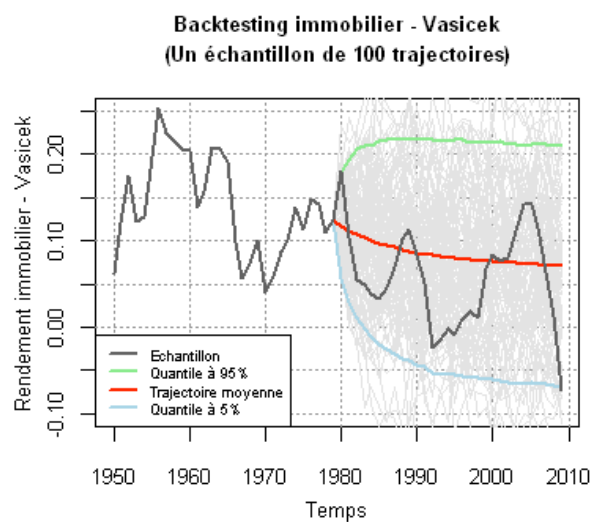
En réalisant plusieurs tests, il s'avère que le modèle de Vasicek ne semble pas adapté aux données, les paramètres ne sont pas tous significatifs et les pouvoirs explicatifs insuffisants. Des tests complémentaires avec le modèle de Black & Scholes ont été réalisés pour comparer les résultats.

Modèle de Black & Scholes

Les résultats semblent bien meilleurs avec ce modèle, les paramètres sont significatifs et les hypothèses vérifiées. Pour l'immobilier il est préférable d'utiliser la totalité des données à disposition selon le critère AICc.

	Test 1	Test 2
N	30	59
K	2	2
Log vraisemblance	44,2445	73,4355
Critère AICc	1,1527	0,4825
Critère BIC	-0,3375	-0,7776

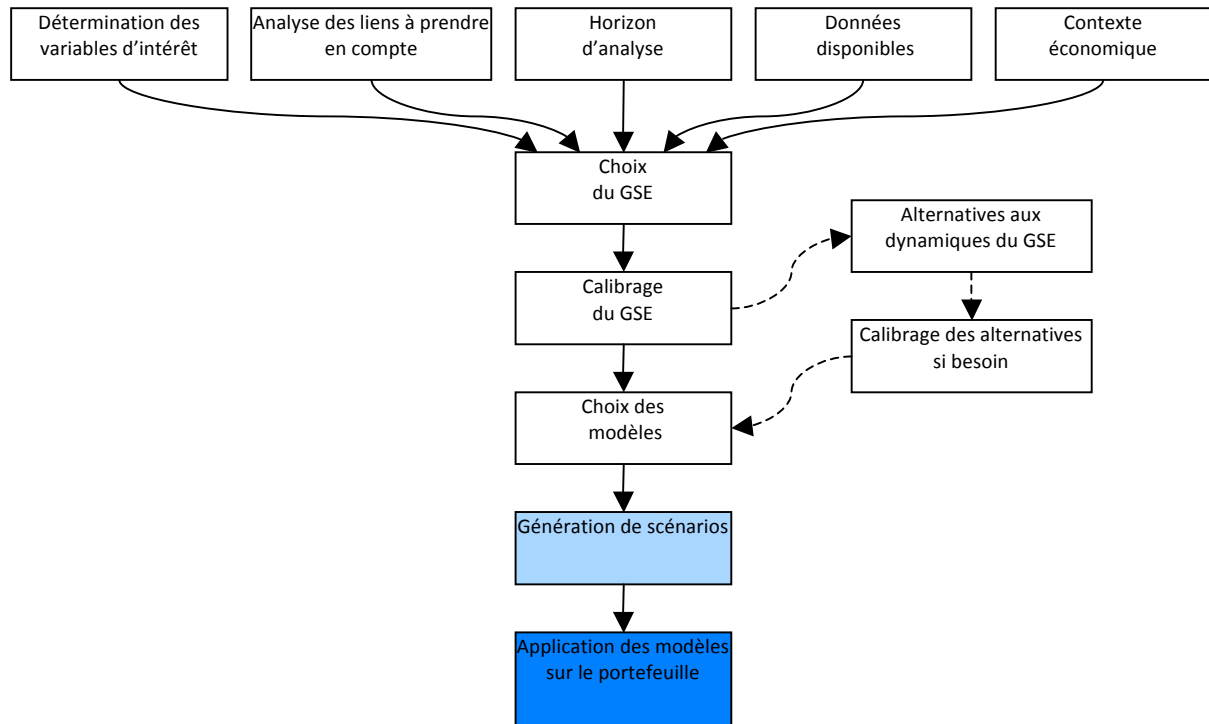
Le backtesting montre que le modèle de B&S ainsi calibré aurait tendance à surestimer les rendements de l'immobilier.



V. Impact du calibrage sur un portefeuille

L'objet de cette partie est de mettre en œuvre l'estimation en valeur de marché d'un portefeuille d'actifs, en se basant sur les éléments détaillés dans les parties précédentes de l'étude.

Après avoir choisi la composition du portefeuille ainsi qu'un GSE, il s'agira de le mettre en place afin de comparer les résultats suivants les modèles et calibrages utilisés.



A. *Le portefeuille et éléments de simulation*

1. Composition du portefeuille

L'allocation de son portefeuille suivant les différentes classes d'actifs va permettre à l'assureur de gérer le risque auquel il s'expose dans ses investissements, c'est pourquoi il est souvent très diversifié afin de minimiser les risques de ruine de l'assureur en cas de krach boursier par exemple.

Les catégories d'actifs venant généralement composer les portefeuilles sont :

1. Les obligations
2. Les actions
3. L'immobilier
4. Le monétaire

La question est alors de déterminer dans quelle proportion chaque type est représenté.

Pour l'étude, la projection est à long terme, or les actifs monétaires servent principalement de trésorerie et sont conservés environ deux ans. La simulation du portefeuille s'intéressera principalement aux obligations, actions et l'immobilier.

En 2009, un document publié conjointement par l'ACAM (devenue l'ACP) et la Banque de France sur le thème « Evolutions récentes de la structure des placements des assureurs », va permettre de choisir un portefeuille fictif.

L'étude se base uniquement sur les entités qui sont du ressort du code des assurances, ceci exclut donc les institutions de prévoyance et les mutuelles.

Il ressort que ce sont les assureurs vie qui investissent principalement dans des actifs de long terme, les assureurs non-vie détenant beaucoup de titres pour une année ou deux seulement, en regard des risques qu'ils couvrent.

Par ailleurs, la durée résiduelle moyenne des titres de créance peut être considérée de 7/8 ans, sachant que certains placements ont une durée résiduelle qui peut aller jusqu'à 100 ans.

Une nouvelle étude publiée en mars 2011, par la banque de France et l'ACP, intitulée « Solvabilité 2 : principaux enseignements de la cinquième étude quantitative d'impact (QIS5) », permet d'avoir une vision globale du portefeuille moyen des assureurs.

Il en ressort que la composition du portefeuille « standard » retenu est :

- 60% pour les obligations, environ ;
- 10% pour l'immobilier ;
- 30% pour les actions.

Pour mieux étudier l'impact d'un titre dont la valeur varie beaucoup, nous considérons que les 30% d'actions sont divisés en :

- 25% d'actions diversifiées ;
- 5% d'actions AXA.

Nous posons quelques hypothèses fortes pour la suite de l'étude :

- La simulation du portefeuille suit la stratégie « constant mix », la réallocation des actifs de façon annuelle permet de conserver la même proportion pour chacun d'eux ;
- La projection est faite sur 30 ans ;
- La dépendance entre classes d'actifs est modélisée par une matrice de corrélation ;
- Les obligations sont de duration 7 ans tous les débuts d'année quand on réalloue le portefeuille.

2. GSE retenu

L'étude menée en seconde partie du document sur les GSE de Wilkie et Alghrim, a montré qu'il était préférable d'utiliser celui d'Alghrim.

3. Matrice de corrélation

L'évolution de chacun des actifs n'est pas indépendante, il est donc nécessaire de tenir compte des corrélations des différentes variables. Pour cela, on s'appuie sur les résidus des calibrages des différents modèles.

La matrice doit comprendre les résidus des variables suivantes :

- Inflation (Inf)
- Taux court (Tx C)
- Taux long (Tx L)
- Rendement action Friggit (Act F)
- Rendement action AXA (Act A)
- Rendement immobilier (Immo)

4. Construction des variables aléatoires

Discretisation du temps

Les données ont permis de calibrer les processus avec un pas de temps annuel, il semble donc logique de conserver cette mesure. La projection sera réalisée sur 30 ans. Le pas de temps est donc défini comme égal à 1.

Remarque : le titre d'AXA a été calibré sur des données mensuelles, ceci n'empêche pas de réaliser des simulations avec un pas annuel.

Simulation de variables aléatoires corrélées

La décomposition de Cholesky nous permet de pouvoir simuler des variables aléatoires corrélées selon un facteur ρ , le facteur de corrélation étant constant.

Le théorème indique qu'une matrice A semi-définie positive, peut s'écrire comme la multiplication de matrice triangulaire inférieure L : $A = LL^T$.

Par exemple : pour avoir deux variables corrélées, la matrice de corrélation A est définie

comme : $A = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ et $L = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}$.

Pour obtenir les variables aléatoires corrélées, il faut d'abord simuler deux variables aléatoires indépendantes U_1 et V . Puis on déduit le couple de variables corrélées (U_1, U_2) :

$$\begin{cases} U_1(t) \\ U_2(t) = \rho U_1(t) + \sqrt{1-\rho^2} V(t) \end{cases}$$

Pour l'étude, la décomposition de Cholesky va nous permettre d'obtenir six variables aléatoires corrélées selon les résidus obtenus lors du calibrage des modèles.

Remarque : pour le modèle de Merton, seul le brownien standard est généré suivant une corrélation, le mouvement de poisson associé à sa loi normale est indépendant.

B. Comparaison de deux calibrages

Il s'agit ici de mettre en regard les performances estimées à long terme pour un même portefeuille, mais avec un choix de calibrage différent. La première vision sera dite « naïve » dans le sens où on retiendra les modèles proposés dans le GSE d'Alghrim et un calibrage sur les données à 30 ans. La seconde vision, dite « optimale » consistera à appliquer ce qui a été retenu comme le plus performant dans la partie précédente.

La première étape pour réaliser la projection des classes d'actifs est de simuler des variables aléatoires respectant la matrice de corrélation. Ensuite une estimation de chaque actif sera possible.

Impact du calibrage sur un portefeuille

Pour l'étude, on choisit de réaliser 10 000 simulations. Pour chacune des simulations, le point de départ au temps t_0 sera la dernière valeur observée dans nos échantillons, sauf lorsque le dernier point était trop particulier.

Portefeuille avec calibrage dit « naïf »

Le calibrage du portefeuille va suivre les éléments présentés dans le tableau :

Variable économique	Modèle	Période de calibrage
Inflation	Vasicek à un facteur	31 données
Taux d'intérêt	Vasicek à deux facteurs	31 données
Action Friggit	Black et scholes	31 données
Action AXA	Black et scholes	Série complète (257 données)
Immobilier	Vasicek à un facteur	31 données

A partir des résidus des modèles ainsi calibrés, on obtient la matrice de corrélation des actifs qui nous permettra de simuler les variables aléatoires selon la décomposition de Cholesky :

	Inf	Tx C	Tx L	Act F	Act A	Immo
Inf	1	-0,38	-0,09	0,07	-0,037	0,39
Tx C	-0,38	1	0,48	0,22	0,10	-0,41
Tx L	-0,09	0,48	1	0,28	-0,38	-0,39
Act F	0,07	0,22	0,28	1	-0,13	0,18
Act A	-0,037	0,1	-0,38	-0,13	1	0,08
Immo	0,39	-0,41	-0,39	-0,18	0,08	1

Remarque : la matrice de corrélation est calculée sur les 30 données de résidus par modèle, sauf pour les coefficients mettant en jeu l'action AXA, les résidus ont été agrégés pour se ramener à une granularité annuelle, la série complète est de 21 résidus.

Portefeuille calibrage dit « optimal »

Le calibrage du portefeuille va suivre les éléments présentés dans le tableau :

Variable économique	Modèle	Période de calibrage
Inflation	Vasicek à un facteur	Série complète
Taux d'intérêt	Vasicek à deux facteurs	Série complète
Action Friggit	Black et scholes	31 données
Action AXA	Merton	Série complète
Immobilier	Black et scholes	Série complète

La matrice de corrélation est donnée par :

	Inf	Tx C	Tx L	Act F	Act A	Immo
Inf	1	-0,85	-0,75	0,06	0,18	0,10
Tx C	-0,85	1	0,87	0,29	-0,16	-0,17
Tx L	-0,75	0,87	1	0,30	-0,39	-0,06
Act F	0,06	0,29	0,30	1	-0,43	0,06
Act A	0,18	-0,16	-0,39	-0,43	1	-0,28
Immo	0,10	-0,17	-0,06	0,06	-0,28	1

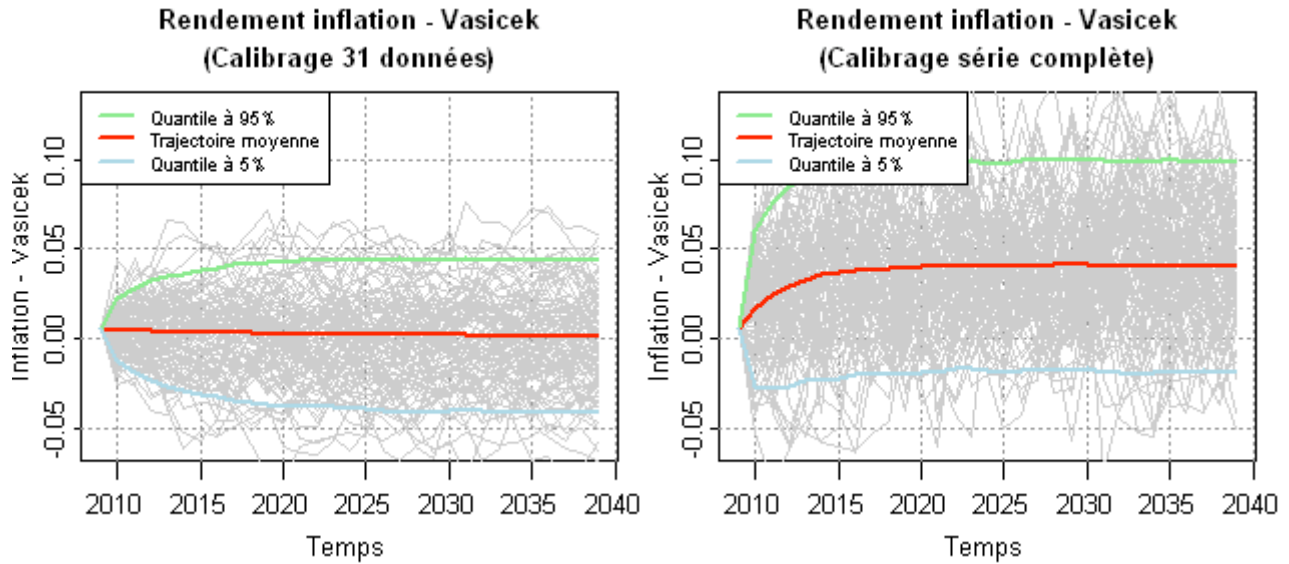
1. Le rendement de l'inflation

Pour effectuer les projections du rendement de l'inflation, il faut utiliser la discrétisation exacte des modèles, soit pour Vasicek :

$$q_{t_{i+1}}^{Rdt} = \exp(-a\delta)q_{t_i}^{Rdt} + \frac{b}{a}[1 - \exp(-a\delta)] + \sigma\sqrt{\frac{1 - \exp[-2a\delta]}{2a}}\varepsilon_{t_i}$$

Il est alors possible d'en déduire le taux d'inflation réel :

$$q_{t_{i+1}}^{réel} = \exp(q_{t_{i+1}}^{Rdt}) * q_t^{réel}$$



Les résultats montrent que l'inflation sera plus importante en moyenne quand on tient compte de la série complète, il en est de même de la volatilité.

L'inflation est faible dans le premier cas et un peu trop forte dans le second, il peut sembler opportun de modifier la valeur de certains paramètres à l'aide d'avis d'experts pour intégrer au mieux le contexte économique.

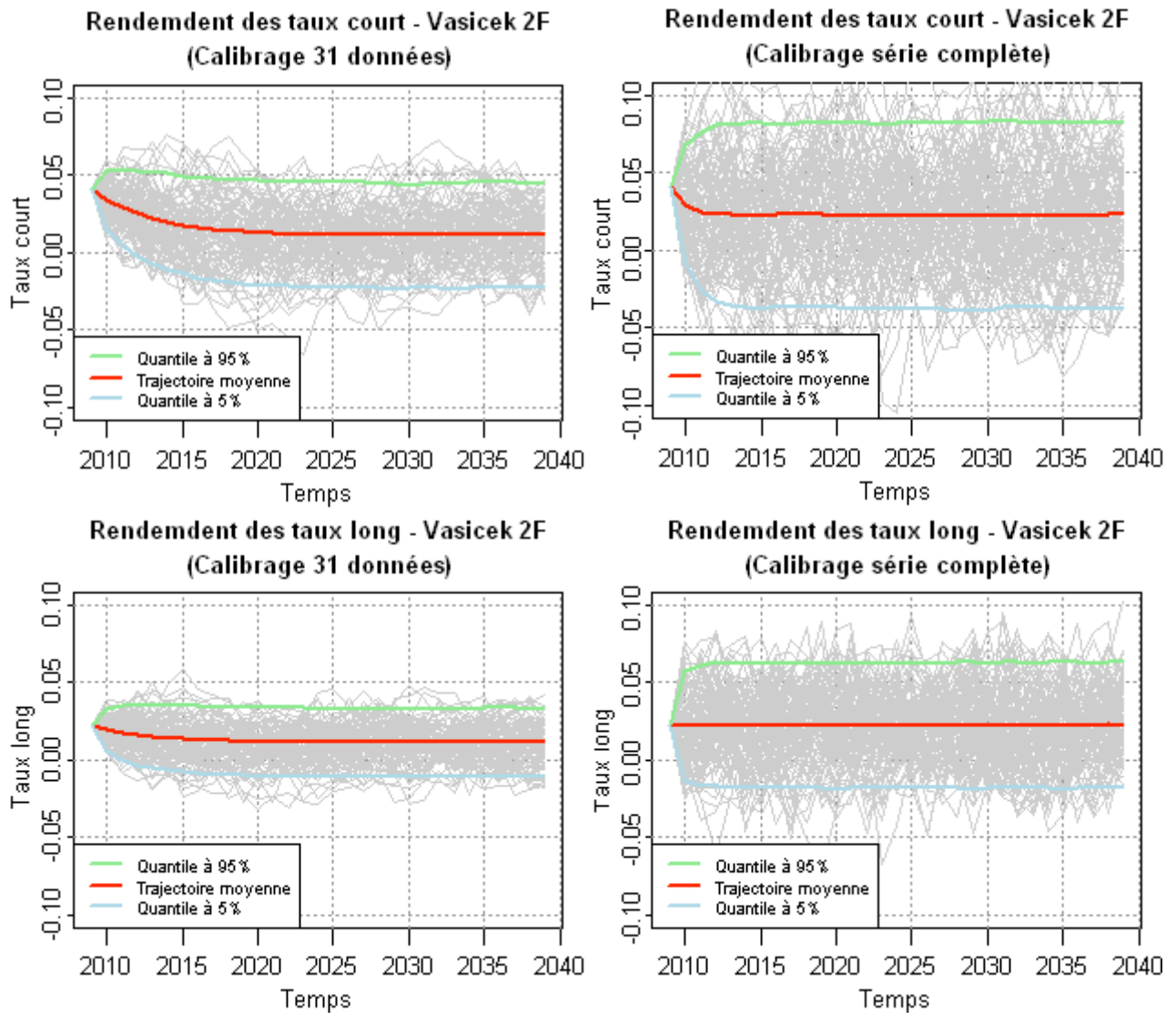
2. Le rendement des taux (court et long)

Taux d'intérêt réels

Pour effectuer les projections, il faut utiliser la discrétisation exacte des modèles, soit pour Vasicek à deux facteurs :

$$r_{t_{i+1}} = \exp(-a_r\delta)r_{t_i} + l_{t_i}[1 - \exp(-a_r\delta)] + \sigma_r\sqrt{\frac{1 - \exp[-2a_r\delta]}{2a_r}}\varepsilon_{r,t_i}$$

$$l_{t_{i+1}} = \exp(-a_l\delta)l_{t_i} + \mu_l[1 - \exp(-a_l\delta)] + \sigma_l\sqrt{\frac{1 - \exp[-2a_l\delta]}{2a_l}}\varepsilon_{l,t_i}$$



La projection des taux est sensiblement la même en moyenne, en revanche la volatilité est bien plus importante dans le cas du calibrage sur la série entière.

La simulation via le modèle de Vasicek peut conduire à des prévisions négatives, c'est le cas ici sur les rendements, certaines trajectoires amènent à avoir des valeurs de taux réels négatifs, ce qui est un inconvénient majeur de ce modèle.

A partir de la projection des rendements, on déduit les taux d'intérêt réels :

$$r_{t+1}^{reel} = (1 + r_{t_i}^{Rdt}) * (1 + q_t^{rdt}) - 1$$

$$r_{t+1}^{reel} = (1 + r_{t_i}^{Rdt}) * (1 + q_t^{rdt}) - 1$$

3. Le prix des obligations

Le prix des obligations est déterminé à partir des variables calibrées et simulées précédemment.

D'après la modélisation d'Ahlgrim qui s'appuie sur les travaux de Hilbert, Mowbray et Turnbull réalisés en 2001, le prix est dépendant du taux court réel, du taux long réel et de l'inflation. Le prix réel de l'obligation en t valant 1 en T est donné par la relation :

$$P_{nom}(t, T) = P_{reel}(t, T) P_{inf}(t, T)$$

Avec :

$$P_{reel}(t, T) = \exp(A_r(t, T) - r_t B_r(t, T) - l_t B_l(t, T))$$

Pour simplifier les expressions, on pose : $s = T - t$

$$B_r(t, T) = \frac{1 - \exp(-a_r s)}{a_r}$$

$$B_l(t, T) = \frac{a_r}{a_r - a_l} \left(\frac{1 - \exp(-a_l s)}{a_l} - \frac{1 - \exp(-a_r s)}{a_r} \right)$$

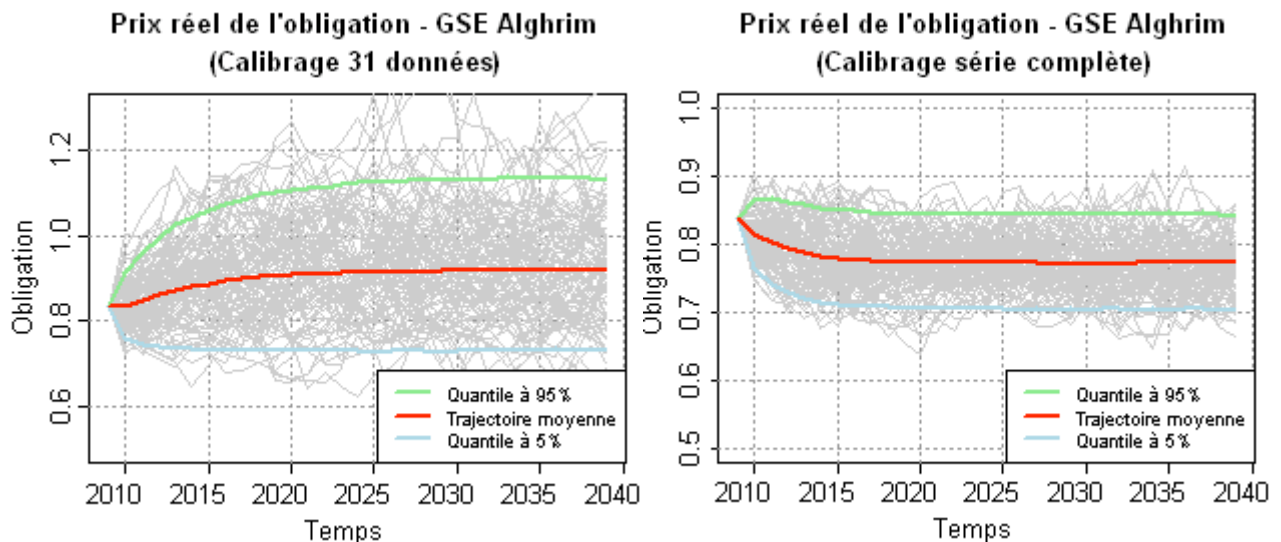
$$A_r(t, T) = (B_r(t, T) - s) \left(\mu - \frac{\sigma_r^2}{2a_r^2} \right) + B_l(t, T) \mu - \frac{\sigma_r^2 B_r(t, T)^2}{4a_r} + \frac{\sigma_l^2}{2} \left[\frac{s}{a_l^2} - 2 \frac{B_l(t, T) + B_r(t, T)}{a_l^2} + \frac{1 - \exp(-2a_r s)}{2a_r(a_r - a_l)^2} - \frac{2a_r(1 - \exp(-(a_r + a_l)s))}{a_l(a_r - a_l)^2(a_r + a_l)} + \frac{a_r^2(1 - \exp(-2a_l s))}{2a_l^3(a_r - a_l)^2} \right]$$

$$P_{inf}(t, T) = A_q(t, T) \exp(-q_t B_q(t, T))$$

$$B_q(t, T) = \frac{1 - \exp(-a_q s)}{a_q}$$

$$A_q(t, T) = \exp \left(\frac{(B_q(t, T) - s)(a_q^2 \mu_q - \sigma_q^2 / 2)}{a_q^2} - \frac{\sigma_q^2 B_q(t, T)^2}{4a_q} \right)$$

Les valeurs des paramètres étant ceux émanant du calibrage du modèle de Vasicek à deux facteurs.



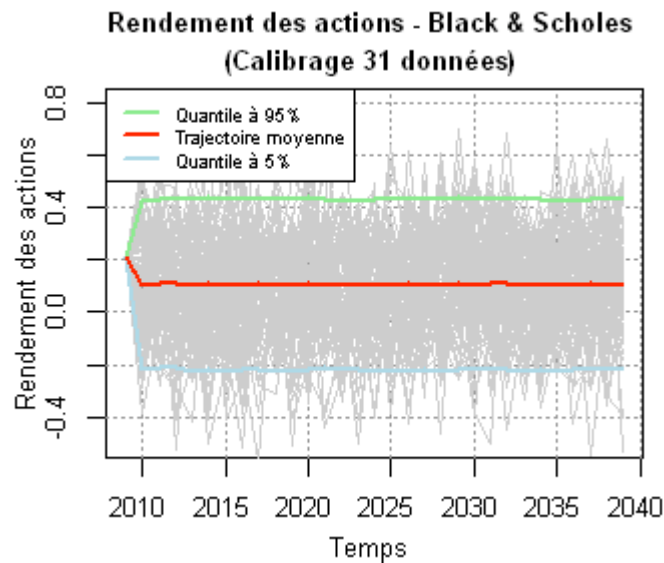
Les obligations sont plus volatiles dans le premier calibrage, la valeur de l'obligation donnant 1 euro à terme est plus élevée aussi.

4. Le rendement des actions Friggit

Les rendements logarithmiques des actions sont obtenus via la formule :

$$r_{ti+1} = \left(\mu - \frac{\sigma^2}{2} \right) \delta + \sigma \varepsilon_{t_i}$$

Dans les deux propositions de calibrage, on retient la même option :



5. Le rendement de l'action AXA

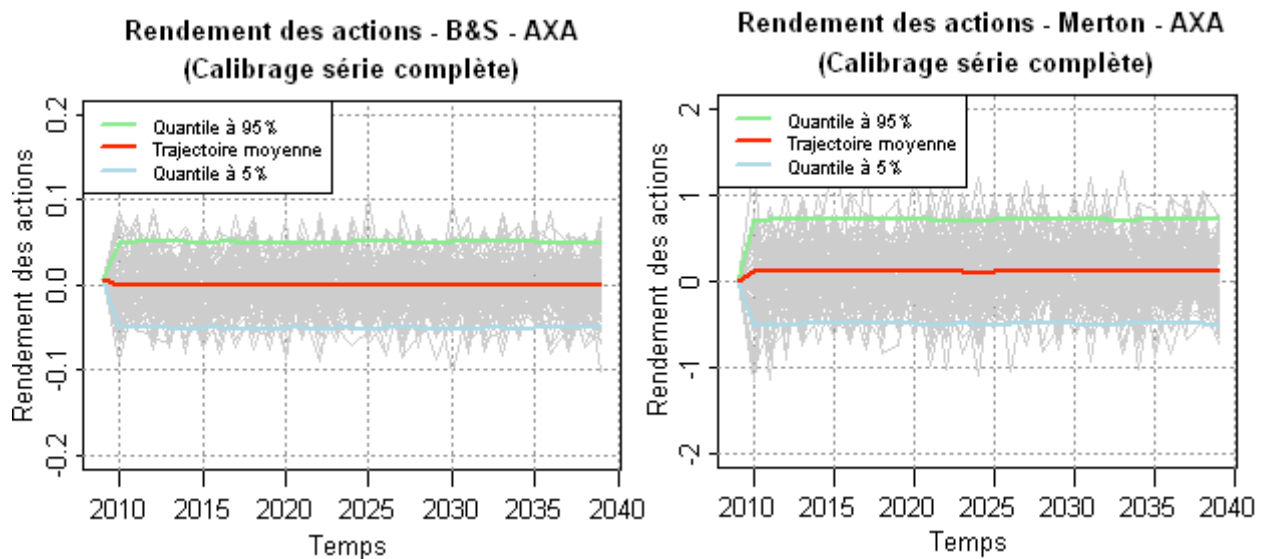
Le rendement de l'action AXA a été modélisé par deux modèles distincts, dans le premier cas, la projection des rendements s'obtient par :

$$r_{t_{i+1}} = \left(\mu - \frac{\sigma^2}{2}\right)\delta + \sigma\epsilon_{t_i}$$

Pour le modèle de Merton, il faut constituer les sauts, puis déduire le rendement :

$$r_{t_i} = \left(\left(\mu - \frac{\sigma^2}{2} \right) \delta + \sigma(W(\delta)) + \sum_{k=N(t_{i-1})}^{N(t_i)} (U(k)) \right)$$

Le calibrage de l'action AXA s'est fait sur des données mensuelles, donc avec le delta=1/12. La valeur des paramètres peut donc être utilisée pour une projection annuelle en utilisant delta=1.



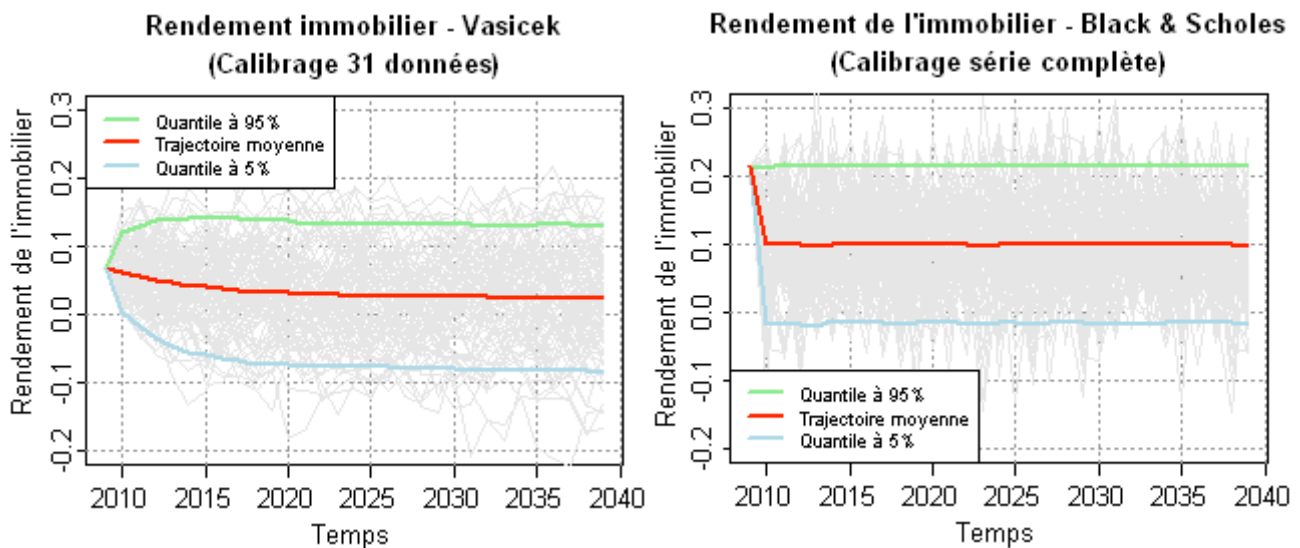
Les résultats obtenus sont très différents, l'échelle des graphiques n'étant pas la même, cela devrait avoir un fort impact sur le portefeuille final.

Le modèle de Merton permet de se rapprocher de données de marché par les effets de saut. Les sauts sont supposés symétriques et en moyenne nuls ce qui intuitivement n'est pas conforme à la réalité.

Le calibrage indique qu'il y aurait en moyenne 2,8 sauts par année, peut être que ce paramètre est un peu fort, si on observe la série de rendement, on s'aperçoit que le modèle comporte beaucoup de sauts à partir de 2006, ceci a dû fortement influencer les paramètres de saut lambda plus particulièrement. Cet exemple met en avant l'importance de l'étude préalable de chacune des séries de données pour exclure les points aberrants des séries (étape non traitée dans ce document).

6. Le rendement de l'immobilier

Les rendements de l'immobilier sont eux aussi simulés par des modèles distincts et des périodes de calibrage distinctes de surcroît. La volatilité est plus grande dans le second cas, tout comme la moyenne de long terme.



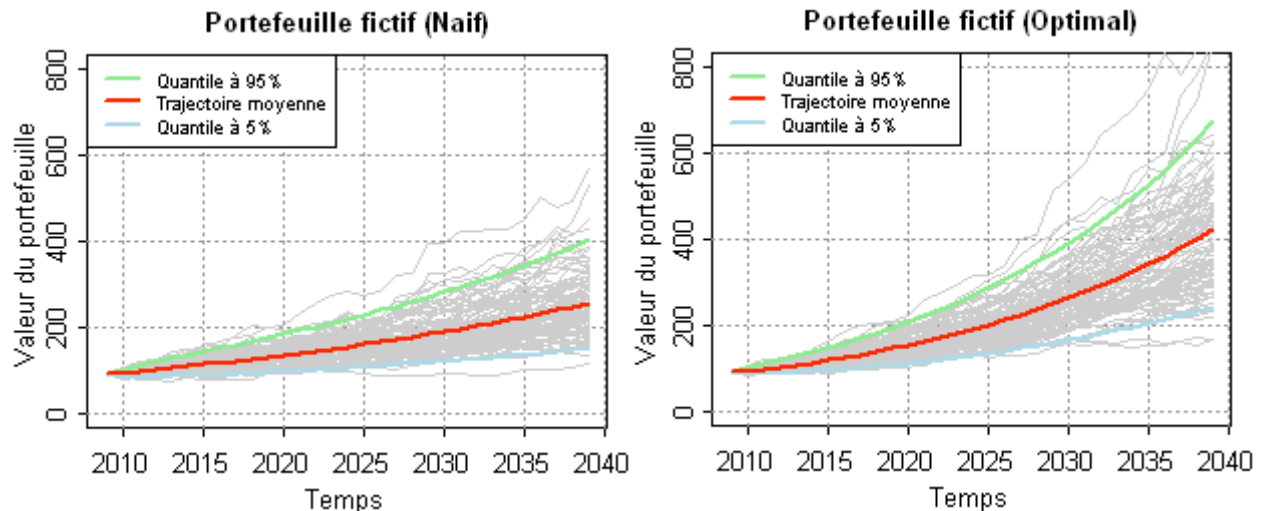
C. Impact du calibrage sur le portefeuille global

La simulation du portefeuille $P(t)$ est établie selon la stratégie « constant mix », ce qui signifie qu'à chaque date t , les proportions du portefeuille sont conservées.

La réallocation du portefeuille est faite à chaque période selon la formule :

$$P(t) = 60\%P(t-1) \frac{B(t)}{B(t-1)} + 25\%P(t-1) \frac{S(t)}{S(t-1)} + 5\%P(t-1) \frac{Axa(t)}{Axa(t-1)} + 10\%P(t-1) \frac{I(t)}{I(t-1)}$$

La valeur initiale des titres d'actions et de l'immobilier est 100, en revanche pour les obligations, la projection donne le prix aujourd'hui d'une obligation valant 1 à maturité, donc la valeur initiale est différente de 100.



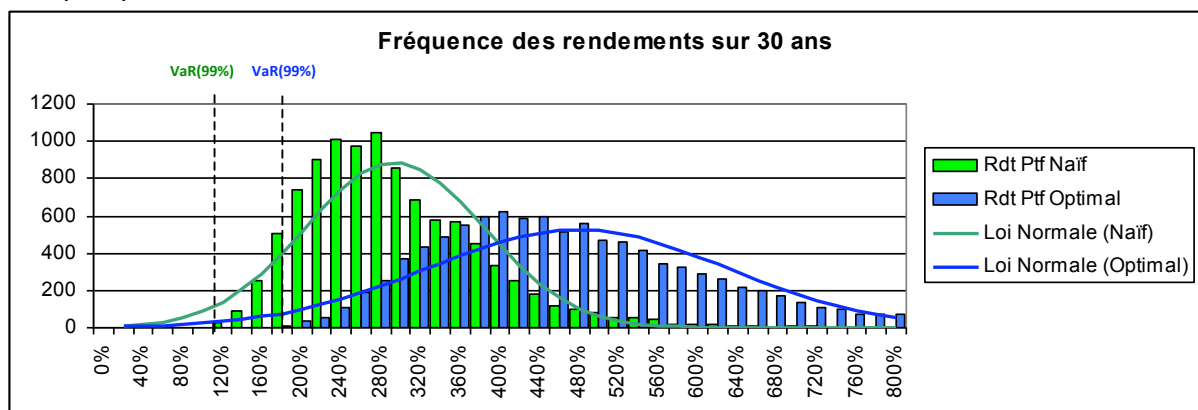
Comme on l'a constaté dans les différentes simulations, la volatilité est plus importante dans le cas du calibrage optimal, ceci s'explique par :

- Une volatilité plus importante lors d'un calibrage sur la série de données complète
- Le choix des modèles, et plus particulièrement le choix du modèle de Merton pour l'action AXA.

En effet, le choix du modèle de Merton influence très fortement la valeur finale du portefeuille. Outre le fait que la valeur des paramètres estimés soit un peu élevée, l'effet est décuplé quand on passe au prix de l'action car ce sont les rendements logarithmiques qui ont été modélisés, ceci signifie que le prix s'obtient en utilisant une exponentielle qui fait augmenter rapidement les valeurs.

Etude des rendements du portefeuille sur 30 ans

En étudiant le graphique de la fréquence des rendements obtenus selon le calibrage, on constate effectivement que le profil est très différent, la probabilité de ruine ou tout autre calcul prospectif seront nécessairement différents dans l'un ou l'autre des cas.



La Value at Risk (VaR) tente de mesurer le risque auquel on s'expose avec le portefeuille détenu à un horizon donné. Il existe différentes méthodes de calcul pour la VaR, par exemple basés sur des valeurs historiques, ou sur des simulations, comme le présente Roland Portait dans « Finance de marché - Instruments de base, produits dérivés, portefeuilles et risques ».

Nous retenons ici, le calcul simple qui consiste à prendre le quantile à 99% sur les rendements obtenus par les 10 000 simulations entre la valeur initiale du portefeuille et sa valeur finale. Les résultats obtenus sont les suivants :

- Portefeuille naïf : 137% de rendement
- Portefeuille optimal : 217% de rendement

La VaR à 99% s'interprète de la façon suivante : dans 1% des cas, la perte sera supérieure ou égale au montant de la VaR.

Dans notre étude, il n'y a pas vraiment de perte, mais plutôt un moindre gain, au bout de 30 ans, dans 1% des cas, le gain sera inférieur à la valeur de la VaR.

Les résultats obtenus dans l'étude sont à modérer dans le sens où un assureur ne gardera jamais un portefeuille en l'état, il peut avoir besoin de revendre une partie de son portefeuille pour faire face à ses engagements.

VI. Conclusion

Après avoir montré que l'évaluation en valeur de marché des actifs détenus par une société d'assurance était devenue incontournable pour la mise en conformité avec les nouvelles normes (Solvabilité II, IFRS notamment), nous avons proposé une démarche complète pour répondre à cette demande.

Nous avons présenté beaucoup d'outils qui sont à notre disposition pour évaluer les actifs. Les générateurs de scénarios économiques semblent tout à fait adaptés, ils permettent au travers de différents processus stochastiques de simuler des grandeurs économiques, tout en tenant compte de leurs interactions. Toutefois, il ne faut pas hésiter à remettre en cause les modèles intégrés au GSE lorsque ceux-ci ne donnent pas de résultats satisfaisants du point de vue du calibrage. Les nombreux modèles existant dans la littérature peuvent permettre de pallier aux défauts rencontrés sur l'un ou plusieurs des modèles du GSE par rapport aux caractéristiques des titres d'actifs à modéliser. Ainsi dans notre étude, pour l'action AXA, le modèle de Merton s'est révélé plus adapté que celui de Black et Scholes car l'action possède des sauts dans sa trajectoire.

Le choix d'un autre modèle peut aussi être induit lorsque le modèle nécessite des données qui sont non observables (ou indisponibles) dans la réalité.

Pour identifier la qualité d'un modèle, il est nécessaire de contrôler que l'ajustement aux données est satisfaisant et que les hypothèses de celui-ci sont vérifiées, elles sont souvent basées sur les propriétés des résidus (suivre une loi normale, caractère iid, etc...). Si plusieurs modèles et/ou calibrages sont bons, il est alors possible de distinguer le meilleur d'entre eux, en utilisant des critères de comparaison tels que : AIC, BIC, LRT.

La mise en œuvre finale de l'étude met en exergue les impacts conséquents en termes de valorisation du portefeuille que peuvent induire les choix de modèle ou de calibrage. Ceux-ci sont d'autant plus grands que la projection est à long terme.

L'enjeu de l'évaluation des actifs en valeur de marché est donc très important en termes de gestion de risques, ou encore pour le calcul du montant de capital de solvabilité requis pour qu'une société d'assurance puisse poursuivre son activité.

Il résulte de l'étude que quels que soient les modèles retenus et la qualité du calibrage réalisé, l'avis d'un expert est essentiel pour réaliser le choix des données et si besoin réajuster les paramètres des modèles, notamment quand la méthode d'estimation a tendance à surestimer les valeurs.

D'autres axes seraient intéressants à développer pour approfondir cette étude :

- Réaliser une étude d'analyse des données pour corriger les valeurs atypiques ou aberrantes en amont des simulations ;
- Etudier de façon plus détaillée la dépendance entre les classes d'actifs, par exemple avec l'étude des copules ;
- Etudier d'autres modèles introduisant une volatilité stochastique comme celui de Hull et White pour la modélisation des taux.

VII. Annexes

A. Modèles AR / MA / ARCH

1. Modèle Auto Régressif d'ordre p (ARp)

Un processus AR admet une mémoire infinie (l'horizon des prévisions n'est donc pas limité). En effet l'équation est la suivante :

$$X_t = c + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t$$

Avec c une constante ;

α_p les paramètres du modèle ;

ε_t perturbation qui suit un bruit blanc.

En utilisant un opérateur de retard ($B^i X_t = X_{t-i}$), le modèle précédent s'écrit :

$$X_t = c + \sum_{i=1}^p (\varphi_i B^i X_t) + \varepsilon_t$$

Ou encore :

$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p) X_t = c + \varepsilon_t$$

2. Modèle Moyenne Mobile d'ordre p (MAp)

Un processus moyenne mobile est en fait une combinaison linéaire de bruits blancs. En effet l'équation est la suivante :

$$Y_t = c + \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2} + \dots + \alpha_p \varepsilon_{t-p} + \varepsilon_t$$

$$Y_t = c + \sum_{i=1}^p (\alpha_i \varepsilon_{t-i}) + \varepsilon_t$$

En utilisant un opérateur de retard ($B^i Y_t = Y_{t-i}$), le modèle précédent s'écrit :

$$Y_t = c + \sum_{i=1}^p (\varphi_i B^i \varepsilon_{t-i}) + \varepsilon_t$$

Ou encore :

$$Y_t = c + (1 + \varphi_1 B + \varphi_2 B^2 + \dots + \varphi_p B^p) \varepsilon_t$$

3. Modèle AutoRegressive Conditional Heteroskedasticity d'ordre p (ARCHp)

Un processus ARCH est utilisé lorsque l'on souhaite prendre en compte la modélisation de la volatilité.

Engle a introduit ce modèle en 1982 afin de tenir compte du fait que l'erreur de prévision peut dépendre pour certaines séries temporelles de la taille de l'erreur des prévisions précédentes.

Une autre motivation pour l'utilisation de ces modèles en finance est la volatilité par classe, cette volatilité décrit la tendance à ce que de fortes variations de prix succèdent à de fortes variations aux périodes précédentes (et inversement).

On a :

$$y_t = X_t' \beta + \varepsilon_t$$

Avec $\varepsilon_t = u_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$ avec α_0 et $\alpha_1 > 0$

$$X_t = \begin{bmatrix} x_{1,t} \\ \dots \\ x_{k,t} \end{bmatrix} \text{ et } \beta = \begin{bmatrix} \beta_1 \\ \dots \\ \beta_k \end{bmatrix}$$

$$u_t \sim N(0,1)$$

La moyenne conditionnelle du terme d'erreur est $E[u_t / u_{t-1}] = 0$.

La variance conditionnelle du terme d'erreur dépend de $\sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$, ε_t n'est plus iid.

B. Algorithme et tests

1. Algorithme de GAUSS - NEWTON

L'algorithme de Gauss-Newton procède par itérations successives, à partir d'un point de départ, on détermine un pas et une direction, on arrive à un nouveau point pour lequel la fonction $SCE(\theta)$ est plus faible, et on reproduit le processus.

A l'itération $i + 1$, on a :

$$\hat{\theta}_{i+1} = \hat{\theta}_i + \Delta$$

Avec $\Delta = \left(J_f(\hat{\theta}_i) J_f(\hat{\theta}_i)' \right)^{-1} J_f(\hat{\theta}_i)' r(\hat{\theta}_i)$ où J_f est le Gradient ou Jacobienne de la fonction f ;

Et $SCE(\hat{\theta}_{i+1}) < SCE(\hat{\theta}_i)$.

Preuve :

Dans le cas d'un modèle linéaire, les équations normales à résoudre ne sont bien souvent pas linéaires en θ .

Les équations sont données, $\forall j = 1 \dots p$, par :

$$\begin{aligned} \text{Min}_{\theta} SCE(\theta) &= \text{Min}_{\theta} \sum_{t=1}^T (y_t - f(\theta, x_t))^2 \\ \Leftrightarrow \frac{\partial}{\partial \theta_j} SCE(\theta) &= 0 \\ \Leftrightarrow \sum_{t=1}^T (y_t - f(\theta, x_t)) \frac{\partial}{\partial \theta_j} f(\theta, x_t) &= 0 \end{aligned}$$

L'algorithme de Gauss-Newton est une possibilité de résoudre numériquement la minimisation.

La fonction est approximée au premier ordre par la formule de Taylor autour de θ^* :

$$f(\theta, x_t) \cong f(\theta^*, x_t) + \sum_{j=1}^p \frac{\partial f(\theta, x_t)}{\partial \theta_j} \bigg|_{\theta=\theta^*} (\theta_j - \theta_j^*) \quad \forall 1 \leq t \leq T$$

Sous forme matricielle : $f(\theta) \cong f(\theta^*) + J_f(\theta^*)(\theta - \theta^*)$

En utilisant cette approximation dans la fonction $SCE(\theta)$, on a :

$$\begin{aligned}
 SCE(\theta) &= (y - f(\theta))(y - f(\theta)) \\
 &\equiv (y - f(\theta^*) - J_f(\theta^*)(\theta - \theta^*))(y - f(\theta^*) - J_f(\theta^*)(\theta - \theta^*)) \\
 &\equiv (y - f(\theta^*))^2 - 2(y - f(\theta^*))J_f(\theta^*)(\theta - \theta^*) + (\theta - \theta^*)^T J_f(\theta^*) J_f(\theta^*)(\theta - \theta^*) \\
 &\equiv SCE(\theta^*) - 2(y - f(\theta^*))J_f(\theta^*)(\theta - \theta^*) + (\theta - \theta^*)^T J_f(\theta^*) J_f(\theta^*)(\theta - \theta^*)
 \end{aligned}$$

On cherche alors à résoudre cette approximation de $SCE(\theta)$ en annulant ses dérivées :

$$\begin{aligned}
 \frac{\partial SCE(\theta)}{\partial \theta} &\equiv \frac{\partial}{\partial \theta} \left(SCE(\theta^*) - 2(y - f(\theta^*))J_f(\theta^*)(\theta - \theta^*) + (\theta - \theta^*)^T J_f(\theta^*) J_f(\theta^*)(\theta - \theta^*) \right) \\
 &\equiv -2J_f(\theta^*)(y - f(\theta^*)) + 2J_f(\theta^*) J_f(\theta^*)(\theta - \theta^*)
 \end{aligned}$$

$$-2J_f(\theta^*)(y - f(\theta^*)) + 2J_f(\theta^*) J_f(\theta^*)(\theta - \theta^*) = 0$$

$$\Rightarrow \theta = \theta^* + \left(J_f(\theta^*) J_f(\theta^*) \right)^{-1} J_f(\theta^*)(y - f(\theta^*))$$

L'algorithme s'appuie directement sur cette estimation :

$$\hat{\theta}_{i+1} = \hat{\theta}_i + \Delta$$

$$\text{Avec } \Delta = \left(J_f(\hat{\theta}_i) J_f(\hat{\theta}_i) \right)^{-1} J_f(\hat{\theta}_i)(y - f(\hat{\theta}_i))$$

2. Principe des tests d'hypothèse

L'idée du test est de poser une hypothèse de travail (H_0) et de comparer les prédictions avec les observations pour conclure si l'hypothèse retenue est acceptable ou non.

Un test d'hypothèse se déroule globalement en cinq étapes :

- 1) Formulation de l'hypothèse nulle H_0 et de H_1 l'hypothèse alternative (négation de H_0), et donc détermination de la nature du test (unilatéral ou bilatéral)
- 2) Choix du seuil de significativité, généralement positionné à 5%. Il représente le risque d'erreur qui peut être de deux types :
 - a. Rejeter H_0 alors qu'elle est vraie
 - b. Accepter H_0 alors qu'elle est fautive
- 3) Déterminer la distribution de référence pour effectuer le test
- 4) Réaliser les calculs
- 5) Conclure le test

A noter que :

- Plus la taille de l'échantillon est grand, plus l'erreur sera petite et le test puissant
- Le test est d'autant plus précis que le seuil de significativité sera petit
- Un test permet de rejeter ou non l'hypothèse nulle, mais pas de l'accepter d'emblée.

3. Test Fisher Snedecor (inclu dans ANOVA)

La statistique de test est le ratio F. On montre que F suit une loi de Fisher à $(p-1, n-p)$ degrés de liberté :

$$F = \frac{S_{SCE}^2}{S_{SCR}^2} = \frac{SCE / (p-1)}{SCR / (n-p)}$$

Preuve :

Il est possible d'utiliser ce test que si les distributions de(s) échantillon(s) suivent une loi normale.

Propriété

Si n variables aléatoires suivent une loi normale centrée réduite, alors la somme de leurs carrés suit une loi de χ^2 à n degrés de liberté.

Cette propriété implique que :

$\frac{S_{SCE}^2}{\sigma_{SCE}^2}$ suit la loi χ^2 à $(p-1)$ degrés de liberté. En effet, la somme des écarts à la moyenne au carré est une variable centrée, en la divisant par la vraie variance, on obtient une variable centrée réduite.

De la même façon, on a $\frac{S_{SCR}^2}{\sigma_{SCR}^2}$ suit la loi χ^2 à $(n-p)$ degrés de liberté.

Propriété

Le ratio de deux variables aléatoires indépendantes suivant chacune une loi de χ^2 , suit une loi de Fisher.

On considère alors le ratio des variables précédentes : $F = \frac{S_{SCE}^2 / \sigma_{SCE}^2}{S_{SCR}^2 / \sigma_{SCR}^2}$ est distribué suivant la loi de Fisher à $(p-1, n-p)$ degrés de liberté.

En se plaçant sous l'hypothèse $H_0 : \sigma_{SCE} = \sigma_{SCR}$, on obtient : $F = \frac{S_{SCE}^2}{S_{SCR}^2}$ distribué suivant la loi de Fisher à $(p-1, n-p)$ degrés de liberté.

C. Données

1. Données de Friggit

Tableau des données de Friggit utilisé pour l'étude avant retraitement pour rendre les données utilisables.

Année	Action	Inflation	Taux d'intérêt à long terme	Taux d'intérêt à court terme	Immobilier
1950	0,00159	0,06955	0,0701	0,025	0,00553
1951	0,00206	0,08087	0,0702	0,027	0,00624
1952	0,00283	0,09048	0,06653	0,038	0,00743
1953	0,00328	0,08896	0,06428	0,0405	0,0084
1954	0,00468	0,08934	0,06385	0,036	0,00955
1955	0,00688	0,09017	0,05998	0,0316	0,01162
1956	0,00724	0,09396	0,06294	0,032	0,01496
1957	0,00946	0,0968	0,06457	0,0535	0,01873
1958	0,00874	0,11142	0,06658	0,059	0,02324
1959	0,01122	0,11826	0,05919	0,0407	0,02856
1960	0,01384	0,12257	0,05798	0,0408	0,035
1961	0,01659	0,1266	0,05642	0,0365	0,04024
1962	0,01832	0,13269	0,05657	0,0361	0,04703
1963	0,01702	0,1392	0,05341	0,0398	0,05794
1964	0,01496	0,14357	0,05523	0,047	0,07116

Année	Action	Inflation	Taux d'intérêt à long terme	Taux d'intérêt à court terme	Immobilier
1965	0,01409	0,14755	0,06217	0,0418	0,08614
1966	0,01392	0,15168	0,06583	0,0456	0,09502
1967	0,01307	0,1558	0,06798	0,0477	0,10048
1968	0,01456	0,16268	0,07155	0,0621	0,10806
1969	0,01858	0,17322	0,08255	0,0897	0,11936
1970	0,01981	0,18238	0,08826	0,0867	0,12424
1971	0,0199	0,19246	0,08549	0,0584	0,13165
1972	0,0229	0,20438	0,08103	0,0496	0,14359
1973	0,02553	0,21904	0,08862	0,0871	0,1588
1974	0,02157	0,24928	0,10913	0,129	0,1823
1975	0,02306	0,27861	0,10489	0,079	0,20405
1976	0,02424	0,30565	0,10446	0,0856	0,23645
1977	0,0213	0,33406	0,10996	0,0907	0,27246
1978	0,02832	0,3643	0,10657	0,0797	0,30424
1979	0,03699	0,40371	0,10633	0,0904	0,34418
1980	0,04273	0,45824	0,13925	0,1184	0,41242
1981	0,03966	0,51965	0,16408	0,153	0,45781
1982	0,04191	0,58105	0,1662	0,1487	0,4835
1983	0,05808	0,63696	0,1462	0,1255	0,5082
1984	0,08271	0,68416	0,13335	0,1174	0,52803
1985	0,10885	0,72402	0,11733	0,0994	0,54545
1986	0,17666	0,74327	0,08775	0,0774	0,57078
1987	0,20288	0,76664	0,09573	0,0798	0,61136
1988	0,17949	0,78726	0,09079	0,0752	0,67735
1989	0,25034	0,81567	0,08788	0,0907	0,75904
1990	0,25396	0,84317	0,09944	0,0996	0,82591
1991	0,25215	0,87015	0,09047	0,0949	0,86889
1992	0,25969	0,89123	0,08598	0,1035	0,84831
1993	0,28764	0,90978	0,06907	0,1073	0,83608
1994	0,31173	0,92495	0,07351	0,0679	0,83454
1995	0,28175	0,94097	0,0759	0,0635	0,82693
1996	0,33302	0,95952	0,06385	0,0373	0,83419
1997	0,43827	0,97133	0,05627	0,0324	0,84894
1998	0,58232	0,97807	0,04718	0,03356	0,85847
1999	0,72067	0,98337	0,04688	0,02739	0,91924
2000	1	1	0,05452	0,04124	1
2001	0,81872	1,0163	0,05047	0,04393	1,07871
2002	0,65121	1,03578	0,04926	0,03292	1,16816
2003	0,55695	1,05739	0,04183	0,02325	1,30514
2004	0,67954	1,07988	0,04153	0,02045	1,50345
2005	0,82021	1,09896	0,03462	0,02088	1,73345
2006	1,02213	1,11744	0,03856	0,02837	1,94327
2007	1,18704	1,13407	0,04356	0,03862	2,07105
2008	0,91818	1,16596	0,04308	0,03857	2,0966
2009	0,74241	1,16704	0,03696	0,00715	1,94761

2. Données de l'action AXA

Tableau des données de l'action AXA utilisé pour l'étude avant retraitement pour rendre les données utilisables.

Date	Action AXA	Date	Action AXA	Date	Action AXA	Date	Action AXA
août-1987	7,9287	Janv-1993	8,3571	Juin-1998	37,2464	nov-2003	27,7946
sept-1987	8,3477	Févr-1993	9,32	Juil-1998	35,0499	déc-2003	27,2291
oct-1987	8,5665	Mars-1993	9,1515	Août-1998	38,451	janv-2004	25,9544
nov-1987	8,1476	Avr-1993	9,2168	Sept-1998	35,3333	févr-2004	23,7431
déc-1987	7,7099	Mai-1993	10,2073	Oct-1998	33,3021	mars-2004	24,8374
janv-1988	8,5853	Juin-1993	9,6639	Nov-1998	32,8298	avr-2004	28,0325
févr-1988	8,579	juil-1993	9,0621	Déc-1998	29,0508	mai-2004	27,7589
mars-1988	7,8537	Août-1993	10,7163	Janv-1999	30,9403	juin-2004	29,7326
avr-1988	7,7912	Sept-1993	10,9295	Févr-1999	33,7178	juil-2004	29,4102
mai-1988	5,6714	Oct-1993	11,3147	Mars-1999	30,449	août-2004	29,8792
juin-1988	5,2525	Nov-1993	11,0983	Avr-1999	31,2615	sept-2004	32,4294
juil-1988	5,4651	Déc-1993	10,6762	Mai-1999	31,1009	oct-2004	32,2828
août-1988	5,6402	Janv-1994	10,8003	Juin-1999	19,5184	nov-2004	30,0257
sept-1988	6,2905	Févr-1994	10,6336	Juil-1999	21,2283	déc-2004	32,2926
oct-1988	6,1279	Mars-1994	10,091	Août-1999	24,7711	janv-2005	31,4914
nov-1988	6,6969	Avr-1994	9,4383	Sept-1999	21,5873	févr-2005	32,0385
déc-1988	7,0221	Mai-1994	10,0555	Oct-1999	20,9449	mars-2005	31,2471
janv-1989	6,9345	Juin-1994	10,1583	Nov-1999	19,3672	avr-2005	28,7262
févr-1989	6,5093	juil-1994	11,1586	Déc-1999	24,5822	mai-2005	29,6349
mars-1989	6,5406	Août-1994	10,6762	Janv-2000	23,8075	juin-2005	29,3906
avr-1989	6,4724	Sept-1994	11,226	Févr-2000	22,0597	juil-2005	25,6777
mai-1989	6,0941	Oct-1994	12,3787	Mars-2000	17,4399	août-2005	25,8438
juin-1989	6,555	Nov-1994	13,1058	Avr-2000	12,8201	sept-2005	21,6131
juil-1989	6,3074	Déc-1994	12,932	Mai-2000	13,8877	oct-2005	21,5056
août-1989	5,7433	Janv-1995	12,6518	Juin-2000	9,8725	nov-2005	20,2549
sept-1989	4,8148	Févr-1995	13,2122	Juil-2000	12,9902	déc-2005	24,0558
oct-1989	6,2936	Mars-1995	12,8256	Août-2000	13,5854	janv-2006	23,108
nov-1989	6,2798	Avr-1995	13,5315	Sept-2000	12,2344	févr-2006	20,4992
déc-1989	6,4036	Mai-1995	14,5423	Oct-2000	11,4503	mars-2006	18,7355
janv-1990	6,4656	Juin-1995	14,0174	Nov-2000	11,3558	avr-2006	19,8592
févr-1990	6,9745	juil-1995	14,3792	Déc-2000	11,6392	mai-2006	24,0851
mars-1990	6,3693	août-1995	14,5884	Janv-2001	11,9226	juin-2006	16,6055
avr-1990	5,2756	sept-1995	16,298	Févr-2001	13,2075	juil-2006	12,5604
mai-1990	5,1243	oct-1995	16,6456	Mars-2001	13,4909	août-2006	16,2
juin-1990	5,4888	nov-1995	19,8982	Avr-2001	15,1631	sept-2006	12,194
juil-1990	5,0349	déc-1995	22,7712	Mai-2001	15,9094	oct-2006	9,9125
août-1990	6,1079	janv-1996	25,8924	Juin-2001	15,1624	nov-2006	9,2217
sept-1990	6,7544	févr-1996	24,7929	Juil-2001	15,0474	déc-2006	10,831
oct-1990	7,7724	mars-1996	22,6293	Août-2001	16,025	janv-2007	12,6386
nov-1990	7,965	avr-1996	29,4393	Sept-2001	18,3157	févr-2007	13,4007
déc-1990	8,9004	mai-1996	28,4107	Oct-2001	18,0857	mars-2007	13,1955
janv-1991	9,1687	juin-1996	19,7563	Nov-2001	16,1113	avr-2007	15,1252
févr-1991	8,5084	juil-1996	20,9977	Déc-2001	17,4339	mai-2007	16,8351
mars-1991	9,265	août-1996	23,5869	Janv-2002	16,2646	juin-2007	19,0531
avr-1991	9,32	sept-1996	26,7792	Févr-2002	16,7918	juil-2007	16,83
mai-1991	10,4137	oct-1996	29,8738	Mars-2002	16,0154	août-2007	16,265
juin-1991	10,6544	nov-1996	28,1521	Avr-2002	15,8046	sept-2007	16,74
juil-1991	11,1428	déc-1996	30,0599	Mai-2002	16,3988	oct-2007	15,405
août-1991	10,4893	janv-1997	28,3382	Juin-2002	16,255	nov-2007	15,8
sept-1991	10,7988	févr-1997	26,8492	Juil-2002	17,2135	déc-2007	16,775
oct-1991	10,5512	mars-1997	27,8496	Août-2002	17,0314	janv-2008	13,03
nov-1991	10,0285	avr-1997	28,5476	Sept-2002	17,5873	févr-2008	14,04

Date	Action AXA	Date	Action AXA	Date	Action AXA	Date	Action AXA
déc-1991	9,4713	mai-1997	26,6398	Oct-2002	18,8141	mars-2008	12,675
janv-1992	8,9417	juin-1997	26,6398	Nov-2002	19,6287	avr-2008	12,9
févr-1992	9,5401	juil-1997	28,315	Déc-2002	18,5074	mai-2008	13,11
mars-1992	7,8068	août-1997	31,8747	Janv-2003	19,1304	juin-2008	13,67
avr-1992	9,2512	sept-1997	31,8747	Févr-2003	19,4083	juil-2008	12,835
mai-1992	8,6838	oct-1997	30,4555	Mars-2003	21,3923	août-2008	12,78
juin-1992	7,9891	nov-1997	29,7575	Avr-2003	21,7469	sept-2008	14,27
juil-1992	8,4018	déc-1997	31,2232	Mai-2003	21,1718	oct-2008	14,965
août-1992	9,1137	janv-1998	36,3418	Juin-2003	21,6319	nov-2008	14,525
sept-1992	8,5634	févr-1998	34,1524	Juil-2003	24,4784	déc-2008	14,625
oct-1992	7,7346	mars-1998	39,1123	Août-2003	25,4273		
nov-1992	7,8034	avr-1998	37,5535	Sept-2003	25,7148		
déc-1992	8,584	mai-1998	39,7972	Oct-2003	27,7946		

D. Détail des calculs d'estimation

1. Modèle de Vasicek à un facteur

a) discrétisation exacte

La dynamique du processus à discrétiser est : $dr_t = \kappa(\mu - r_t)dt + \sigma dB_t$, où r_t est une variable aléatoire continue.

L'astuce est d'étudier la dérivée de $\exp(\kappa t)r_t$.

$$\text{Soit } d(\exp(\kappa t)r_t) = \underbrace{r_t \kappa \exp(\kappa t) dt}_{\text{dérivée par rapport à } t} + \underbrace{\exp(\kappa t) dr_t}_{\text{dérivée par rapport à } r}$$

$$\text{On a } dr_t = \kappa(\mu - r_t)dt + \sigma dB_t$$

En multipliant chaque membre par $\exp(\kappa t)$, on obtient :

$$\begin{aligned} \exp(\kappa t) dr_t &= \exp(\kappa t) [\kappa(\mu - r_t)dt + \sigma dB_t] \\ &= \kappa\mu \exp(\kappa t)dt + \sigma \exp(\kappa t)dB_t - r_t \kappa \exp(\kappa t)dt \\ \Rightarrow \exp(\kappa t) dr_t + r_t \kappa \exp(\kappa t)dt &= \kappa\mu \exp(\kappa t)dt + \sigma \exp(\kappa t)dB_t \\ \Rightarrow d(\exp(\kappa t)r_t) &= \kappa\mu \exp(\kappa t)dt + \sigma \exp(\kappa t)dB_t \end{aligned}$$

En intégrant l'équation sur un pas allant de t_0 à t , il advient :

$$\begin{aligned} \int_{t_0}^t d(\exp(\kappa s)r_s) &= \int_{t_0}^t \kappa\mu \exp(\kappa s)ds + \int_{t_0}^t \sigma \exp(\kappa s)dB_s \\ [\exp(\kappa t)r_t]_{t_0} &= \kappa\mu \left[\frac{\exp(\kappa s)}{\kappa} \right]_{t_0}^t + \sigma \int_{t_0}^t \exp(\kappa s)dB_s \\ \exp(\kappa t)r_t &= \exp(\kappa t_0)r_{t_0} + \mu[\exp(\kappa(t - t_0))] + \sigma \int_{t_0}^t \exp(\kappa s)dB_s \end{aligned}$$

En simplifiant l'expression par $\exp(\kappa t)$, on obtient :

$$r_t = \frac{\exp(\kappa t_0)}{\exp(\kappa t)} r_{t_0} + \frac{\mu[\exp(\kappa t) - \exp(\kappa t_0)]}{\exp(\kappa t)} + \sigma \int_{t_0}^t \frac{\exp(\kappa s)}{\exp(\kappa t)} dB_s$$

$$r_t = \exp(-\kappa(t - t_0))r_{t_0} + \mu[1 - \exp(-\kappa(t - t_0))] + \sigma \int_{t_0}^t \exp(-\kappa(t - s))dB_s$$

Les propriétés de l'intégrale d'une fonction déterministe par rapport à un mouvement brownien permettent d'écrire :

$$r_t = \exp(-\kappa(t - t_0))r_{t_0} + \mu[1 - \exp(-\kappa(t - t_0))] + \sigma \sqrt{\frac{1 - \exp[-2\kappa(t - t_0)]}{2\kappa}} \varepsilon_t^{t_0}$$

Avec $\varepsilon_t^{t_0}$ une variable aléatoire de loi normale centrée réduite.

b) Estimation des paramètres de Vasicek à un facteur, maximum de vraisemblance

Étape 1 : Trouver la fonction de densité de la variable aléatoire

Le modèle de Vasicek définit le processus du taux d'intérêt comme :

$$r_{t_{i+1}} = \exp(-a\delta)r_{t_i} + \frac{b}{a}[1 - \exp(-a\delta)] + \sigma \int_{t_i}^{t_{i+1}} \exp(-a(t_{i+1} - s))dB_s$$

Comme le relate Brigo et Mercurio dans leur livre, $r_{t_{i+1}}$ est normalement distribué de moyenne et de variance :

$$E[r_{t_{i+1}} / r_{t_i}] = \exp(-a\delta)r_{t_i} + \frac{b}{a}[1 - \exp(-a\delta)]$$

$$Var[r_{t_{i+1}} / r_{t_i}] = \sigma^2 \frac{1 - \exp[-2a\delta]}{2a}$$

Preuve :

Théorème

Soit f tel que $\int_0^T |f(s)|^2 ds < \infty$, $\forall T$ et $M_t = \int_0^t f(s)dB_s$ alors M est un processus gaussien d'espérance nulle et de variance $\int_0^T f^2(s)dB_s$.

En appliquant le théorème précédemment énoncé, on a bien :

$$\sigma \int_{t_i}^{t_{i+1}} |\exp(-a(t_{i+1} - s))|^2 ds < \infty$$

Donc l'espérance de $\sigma \int_{t_i}^{t_{i+1}} \exp(-a(t_{i+1} - s))dB_s$ est nulle, de plus

$$\sigma \int_{t_i}^{t_{i+1}} [\exp(-a(t_{i+1} - s))]^2 dB_s = \sigma^2 \frac{1 - \exp[-2a\delta]}{2a} \varepsilon_{t_i}^{t_i}, \text{ d'où la variance } = \sigma^2 \frac{1 - \exp[-2a\delta]}{2a}.$$

Si on considère l'ensemble de l'équation, il n'y a pas d'incertitude sur la première partie, donc :

$$E[r_{t_{i+1}} / r_{t_i}] = \exp(-a\delta)r_{t_i} + \frac{b}{a}[1 - \exp(-a\delta)]$$

$$Var[r_{t_{i+1}} / r_{t_i}] = \sigma^2 \frac{1 - \exp[-2a\delta]}{2a}$$

Pour que les conditions d'application de la vraisemblance soit réunies, il faut vérifier que $r_{t_{i+1}}$ et $r_{t_{j+1}}$ sont indépendants, $\forall i \neq j$.

Cette dernière condition est vérifiée car $\int_i^{i+1} \exp(-a(t_{i+1} - s)) dB_s$ est indépendant de $\int_j^{j+1} \exp(-a(t_{j+1} - s)) dB_s$ (les accroissements sont indépendants dans un processus).

Propriété

Soit X une variable aléatoire suit une loi log normale d'espérance μ et de variance σ^2 si $Y = \ln(X)$ suit une loi normale d'espérance μ et de variance σ^2 .

Étape 2 : Définir la fonction de log vraisemblance et sa maximisation

$r_{t_{i+1}}$ suit une loi normale, voici un petit rappel :

Propriété

Soit X une variable aléatoire qui suit une loi normale d'espérance μ et de variance σ^2 , alors la densité de probabilité est définie par :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

La fonction de vraisemblance est donnée par :

$$L(\theta) = \prod_{i=0}^{n-1} \frac{1}{\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{r_{t_{i+1}} - \left(\exp(-a\delta) r_{t_i} + \frac{b}{a} [1 - \exp(-a\delta)] \right)}{\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}}} \right)^2 \right)$$

La fonction de log vraisemblance devient alors :

$$\Lambda = -n \log \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \sqrt{2\pi} \right) - \frac{1}{2 \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \right)^2} \sum_{i=0}^{n-1} \left(r_{t_{i+1}} - \exp^{-2a\delta} r_{t_i} - \frac{b}{a} (1 - \exp^{-2a\delta}) \right)^2$$

où $\Lambda = \sum_{i=0}^{n-1} \Lambda_{i+1}$ avec

$$\Lambda_{i+1} = -\log \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \sqrt{2\pi} \right) - \frac{1}{2 \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \sqrt{2\pi} \right)^2} \left(r_{t_{i+1}} - \exp^{-2a\delta} r_{t_i} - \frac{b}{a} (1 - \exp^{-2a\delta}) \right)^2$$

On remarque que :

$$\begin{aligned} \log \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \sqrt{2\pi} \right) &= \log(\sigma) + \frac{1}{2} \log \left(\frac{1 - \exp^{-2a\delta}}{2a} \right) \\ &= \log(\sigma) + \frac{1}{2} \log(1 - \exp^{-2a\delta}) - \frac{1}{2} \log(2a) \end{aligned}$$

Et $\varepsilon'_{i+1} = r_{t_{i+1}} - \exp^{-a\delta} r_{t_i} - \frac{b}{a} (1 - \exp^{-a\delta})$ représente les résidus.

D'où :

$$\Lambda_{i+1} = -\log(\sigma) - \frac{1}{2} \log(1 - \exp^{-2a\delta}) + \frac{1}{2} \log(2a) - \frac{1}{2} \log(2\pi) - \frac{a(r_{t_{i+1}} - r_{t_i} \exp^{-a\delta} - \frac{b}{a}(1 - \exp^{-a\delta}))^2}{\sigma^2(1 - \exp^{-2a\delta})}$$

$$\Lambda_{i+1} = -\log(\sigma) - \frac{1}{2} \log(1 - \exp^{-2a\delta}) + \frac{1}{2} \log(2a) - \frac{1}{2} \log(2\pi) - \frac{a \sigma^2 (1 - \exp^{-2a\delta})}{\sigma^2(1 - \exp^{-2a\delta})} (\varepsilon'_{i+1})^2$$

Pour maximiser la fonction de log vraisemblance, il faut trouver les valeurs de paramètres qui annulent les dérivées premières.

Pour faciliter l'écriture des équations, il est préférable de dériver les résidus puis la fonction de log vraisemblance.

Gradient des résidus : $\varepsilon'_{i+1} = r_{t_{i+1}} - r_{t_i} \exp^{-a\delta} - \frac{b}{a}(1 - \exp^{-a\delta})$

$$\frac{\partial \varepsilon'_{i+1}}{\partial a} = \delta \exp^{-a\delta} r_{t_i} + \frac{b}{a^2}(1 - \exp^{-a\delta}) - \frac{\delta b}{a} \exp^{-a\delta}$$

$$\frac{\partial \varepsilon'_{i+1}}{\partial b} = -\frac{1}{a}(1 - \exp^{-a\delta})$$

$$\frac{\partial \varepsilon'_{i+1}}{\partial \sigma} = 0$$

Dérivées de la fonction log vraisemblance :

$$\frac{\partial \Lambda_{i+1}}{\partial a} = -\frac{\delta \exp^{-2a\delta}}{1 - \exp^{-2a\delta}} + \frac{0.5}{a} - \frac{\varepsilon'^2_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})} + \frac{2a\delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})^2} - \frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})}$$

$$\frac{\partial \Lambda_{i+1}}{\partial b} = -\frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial b} \varepsilon'_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})}$$

$$\frac{\partial \Lambda_{i+1}}{\partial \sigma} = -\frac{1}{\sigma} + \frac{2a \varepsilon'^2_{i+1}}{\sigma^3(1 - \exp^{-2a\delta})}$$

Il n'est pas possible de résoudre le système d'équation formé par
$$\begin{cases} \frac{\partial \Lambda_{i+1}}{\partial a} = 0 \\ \frac{\partial \Lambda_{i+1}}{\partial b} = 0 \\ \frac{\partial \Lambda_{i+1}}{\partial \sigma} = 0 \end{cases}$$
 car il est

composé d'équations non linéaires, on ne peut pas trouver la valeur des estimateurs de façon analytique.

Pour tenter de résoudre le problème, on simplifie la fonction de vraisemblance, en posant :

$$\alpha = \exp^{-a\delta}, \quad \beta = \frac{b}{a}(1 - \alpha) \quad \text{et} \quad v = \sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}}$$

La fonction de vraisemblance est :

$$L(\theta) = \prod_{i=0}^{n-1} \frac{1}{\nu \sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{r_{t_{i+1}} - (\alpha r_{t_i} + \beta[1 - \alpha])}{\nu} \right)^2 \right)$$

La fonction de log vraisemblance devient alors :

$$\Lambda = -n \log(\nu \sqrt{2\pi}) - \frac{1}{2\nu^2} \sum_{i=0}^{n-1} (r_{t_{i+1}} - \alpha r_{t_i} - \beta(1 - \alpha))^2$$

où $\Lambda = \sum_{i=0}^{n-1} \Lambda_{i+1}$ avec $\Lambda_{i+1} = -\log(\nu \sqrt{2\pi}) - \frac{1}{2(\nu \sqrt{2\pi})^2} (r_{t_{i+1}} - \alpha r_{t_i} - \beta(1 - \alpha))^2$

Le gradient des résidus :

$$\begin{aligned} \frac{\partial \varepsilon'_{i+1}}{\partial \alpha} &= -r_{t_i} + \beta \\ \frac{\partial \varepsilon'_{i+1}}{\partial \beta} &= -(1 - \alpha) \\ \frac{\partial \varepsilon'_{i+1}}{\partial \nu} &= 0 \end{aligned}$$

Le gradient de log vraisemblance :

$$\begin{aligned} \frac{\partial \Lambda_{i+1}}{\partial \alpha} &= -\frac{\varepsilon'_{i+1} \frac{\partial \varepsilon'_{i+1}}{\partial \alpha}}{\nu^2} \\ \frac{\partial \Lambda_{i+1}}{\partial \beta} &= -\frac{\varepsilon'_{i+1} \frac{\partial \varepsilon'_{i+1}}{\partial \beta}}{\nu^2} \\ \frac{\partial \Lambda_{i+1}}{\partial \nu} &= -\frac{1}{\nu} + \frac{\varepsilon'^2_{i+1}}{\nu^3} \end{aligned}$$

La vraisemblance sera maximisée pour les paramètres vérifiant le système d'équation :

$$\left\{ \begin{aligned} \sum_{i=0}^{n-1} \left(\frac{\partial \Lambda_{i+1}}{\partial \alpha} \right) &= -\sum_{i=0}^{n-1} \left(\frac{\varepsilon'_{i+1} \frac{\partial \varepsilon'_{i+1}}{\partial \alpha}}{\nu^2} \right) = 0 \\ \sum_{i=0}^{n-1} \left(\frac{\partial \Lambda_{i+1}}{\partial \beta} \right) &= -\sum_{i=0}^{n-1} \left(\frac{\varepsilon'_{i+1} \frac{\partial \varepsilon'_{i+1}}{\partial \beta}}{\nu^2} \right) = 0 \\ \sum_{i=0}^{n-1} \left(\frac{\partial \Lambda_{i+1}}{\partial \nu} \right) &= \sum_{i=0}^{n-1} \left(-\frac{1}{\nu} + \frac{\varepsilon'^2_{i+1}}{\nu^3} \right) = 0 \end{aligned} \right.$$

En observant les différentes équations, la première constatation est que ν est non nul. On déduit sa valeur de la troisième équation du système :

$$\sum_{i=0}^{n-1} \left(-\frac{1}{v} + \frac{\varepsilon'_{i+1}{}^2}{v^3} \right) = -\frac{n}{v} + \frac{1}{v^3} \sum_{i=0}^{n-1} (\varepsilon'_{i+1}{}^2) = 0$$

$$\hat{v}^2 = \frac{\sum_{i=0}^{n-1} (\varepsilon'_{i+1}{}^2)}{n}$$

De la seconde équation, on déduit l'expression de β :

$$-\sum_{i=0}^{n-1} \left(\frac{\varepsilon'_{i+1} \frac{\partial \varepsilon'_{i+1}}{\partial \beta}}{v^2} \right) = -\sum_{i=0}^{n-1} \left(\frac{\varepsilon'_{i+1} (-(1-\alpha))}{v^2} \right) = 0 \Rightarrow (1-\alpha) \sum_{i=0}^{n-1} (\varepsilon'_{i+1}) = 0$$

$$\sum_{i=0}^{n-1} (\varepsilon'_{i+1}) = \sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i}) - n\beta(1-\alpha) = 0$$

$$\hat{\beta} = \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)}$$

Enfin l'expression de α est déduite de la première équation :

$$-\sum_{i=0}^{n-1} \left(\frac{\varepsilon'_{i+1} \frac{\partial \varepsilon'_{i+1}}{\partial \alpha}}{v^2} \right) = -\sum_{i=0}^{n-1} \left(\frac{\varepsilon'_{i+1} (-r_{t_i} + \beta)}{v^2} \right) = 0 \Rightarrow -\sum_{i=0}^{n-1} (\varepsilon'_{i+1} (-r_{t_i} + \beta)) = 0$$

Or $\varepsilon'_{i+1} = r_{t_{i+1}} - \alpha r_{t_i} - \beta(1-\alpha)$

$$-\sum_{i=0}^{n-1} \left(\left(r_{t_{i+1}} - \alpha r_{t_i} - \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)} (1-\alpha) \right) \left(-r_{t_i} + \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)} \right) \right) = 0$$

$$-\sum_{i=0}^{n-1} \left(\left(r_{t_{i+1}} - \alpha r_{t_i} - \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n} \right) \left(-r_{t_i} + \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)} \right) \right) = 0$$

$$-\sum_{i=0}^{n-1} \left(\begin{aligned} & -r_{t_{i+1}} r_{t_i} + r_{t_{i+1}} \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)} + \alpha r_{t_i} r_{t_i} - \alpha r_{t_i} \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)} \\ & + r_{t_i} \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n} - \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n} \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)} \end{aligned} \right) = 0$$

$$\begin{aligned}
& \sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) - \sum_{i=0}^{n-1} (r_{t_{i+1}}) \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)} - \alpha \sum_{i=0}^{n-1} (r_{t_i}^2) + \alpha \sum_{i=0}^{n-1} (r_{t_i}) \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)} \\
& - \sum_{i=0}^{n-1} (r_{t_i}) \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n} + \left(\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i}) \right) \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \alpha \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)} = 0 \\
& \sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) - \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}})^2}{n(1-\alpha)} + \frac{\alpha \sum_{i=0}^{n-1} (r_{t_{i+1}}) \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)} - \alpha \sum_{i=0}^{n-1} (r_{t_i}^2) + \frac{\alpha \sum_{i=0}^{n-1} (r_{t_i}) \sum_{i=0}^{n-1} (r_{t_{i+1}})}{n(1-\alpha)} - \frac{\alpha^2 \sum_{i=0}^{n-1} (r_{t_i})^2}{n(1-\alpha)} \\
& - \frac{\sum_{i=0}^{n-1} (r_{t_i}) \sum_{i=0}^{n-1} (r_{t_{i+1}})}{n} + \frac{\alpha \sum_{i=0}^{n-1} (r_{t_i})^2}{n} + \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}})^2}{n(1-\alpha)} - \frac{\alpha \sum_{i=0}^{n-1} (r_{t_{i+1}}) \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)} - \alpha \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) \sum_{i=0}^{n-1} (r_{t_i})}{n(1-\alpha)} + \alpha^2 \frac{\sum_{i=0}^{n-1} (r_{t_i})^2}{n(1-\alpha)} = 0 \\
& \sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) - \alpha \sum_{i=0}^{n-1} (r_{t_i}^2) - \frac{\sum_{i=0}^{n-1} (r_{t_i}) \sum_{i=0}^{n-1} (r_{t_{i+1}})}{n} + \alpha \frac{\sum_{i=0}^{n-1} (r_{t_i})^2}{n} = 0 \\
& \sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) - \frac{\sum_{i=0}^{n-1} (r_{t_i}) \sum_{i=0}^{n-1} (r_{t_{i+1}})}{n} = \alpha \sum_{i=0}^{n-1} (r_{t_i}^2) - \alpha \frac{\sum_{i=0}^{n-1} (r_{t_i})^2}{n} \\
& \alpha = \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) - \frac{\sum_{i=0}^{n-1} (r_{t_i}) \sum_{i=0}^{n-1} (r_{t_{i+1}})}{n}}{\sum_{i=0}^{n-1} (r_{t_i}^2) - \frac{\sum_{i=0}^{n-1} (r_{t_i})^2}{n}} \\
& \hat{\alpha} = \frac{n \sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) - \sum_{i=0}^{n-1} (r_{t_i}) \sum_{i=0}^{n-1} (r_{t_{i+1}})}{n \sum_{i=0}^{n-1} (r_{t_i}^2) - \sum_{i=0}^{n-1} (r_{t_i})^2}
\end{aligned}$$

A partir de $\hat{\alpha}$ et $\hat{\beta}$, on obtient la valeur estimée des paramètres $\hat{a}$, $\hat{b}$ et $\hat{v}$:

- $\alpha = \exp^{-a\delta} \Rightarrow \hat{a} = -\frac{\ln(\hat{\alpha})}{\delta}$
- $\beta = \frac{b}{a}(1-\alpha) \Rightarrow \hat{b} = \frac{\hat{a}\hat{\beta}}{1-\hat{\alpha}}$
- $v = \sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \Rightarrow \hat{v} = \sigma \sqrt{\frac{1 - \exp^{-2\hat{a}\delta}}{2\hat{a}}} = \sigma \sqrt{\frac{\delta(1 - \exp^{2\ln(\hat{\alpha})})}{-2\ln(\hat{\alpha})}}$

$$\Rightarrow \hat{\sigma} = \sqrt{\frac{2\hat{a}}{1 - \exp^{-2a\delta}} * \frac{\sum_i (\varepsilon_i)^2}{n}}$$

Dans le cadre du maximum de vraisemblance, il n'y a pas de degré de liberté à retirer au n .

Remarque : l'expression des paramètres est équivalente à celle de la méthode de MCO.

La moyenne de la variable r_{t_i} , noté $\bar{r}_{t_i}$, est : $\frac{\sum_{i=0}^{n-1} (r_{t_i})}{n}$.

$$\hat{\alpha} = \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) - \frac{\sum_{i=0}^{n-1} (r_{t_i}) \sum_{i=0}^{n-1} (r_{t_{i+1}})}{n}}{\sum_{i=0}^{n-1} (r_{t_i}^2) - \frac{\sum_{i=0}^{n-1} (r_{t_i})^2}{n}} \quad \text{d'où } \hat{\alpha} = \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) - n \bar{r}_{t_{i+1}} \bar{r}_{t_i}}{\sum_{i=0}^{n-1} (r_{t_i}^2) - n \bar{r}_{t_i}^2} ;$$

$$\hat{\beta} = \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}}) - \hat{\alpha} \sum_{i=0}^{n-1} (r_{t_i})}{n(1 - \hat{\alpha})} \quad \text{d'où } (1 - \hat{\alpha})\hat{\beta} = \bar{r}_{t_{i+1}} - \hat{\alpha} \bar{r}_{t_i} ;$$

Pour faciliter les démarches de calcul, il est opportun d'utiliser les valeurs des estimations de $\hat{\alpha}$, $\hat{\beta}$ de la méthode des moindres carrés ordinaires. Ensuite, on déduira $\hat{v}$, $\hat{a}$ et $\hat{b}$.

Matrice de variance-covariance / corrélation

Comme pour l'estimation des paramètres du processus, le calcul s'effectue en deux temps, pour les résidus puis pour la log vraisemblance.

On rappelle que la fonction de log vraisemblance est donnée par l'expression :

$$\Lambda = -n \log \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \sqrt{2\pi} \right) - \frac{1}{2 \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \right)^2} \sum_{i=0}^{n-1} \left(r_{t_{i+1}} - \exp^{-2a\delta} r_{t_i} - \frac{b}{a} (1 - \exp^{-2a\delta}) \right)^2$$

où $\Lambda = \sum_{i=0}^{n-1} \Lambda_{i+1}$ avec

$$\Lambda_{i+1} = -\log \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \sqrt{2\pi} \right) - \frac{1}{2 \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \right)^2} \left(r_{t_{i+1}} - \exp^{-2a\delta} r_{t_i} - \frac{b}{a} (1 - \exp^{-2a\delta}) \right)^2$$

$$\Lambda_{i+1} = -\log(\sigma) - \frac{1}{2} \log(1 - \exp^{-2a\delta}) + \frac{1}{2} \log \left(\frac{1 - \exp^{-2a\delta}}{2a} \right) + \frac{1}{2 \left(\sigma \sqrt{\frac{1 - \exp^{-2a\delta}}{2a}} \right)^2} \left(\varepsilon'_{i+1} \right)^2$$

Gradient des résidus : $\varepsilon'_{i+1} = r_{t_{i+1}} - r_{t_i} \exp^{-a\delta} - \frac{b}{a} (1 - \exp^{-a\delta})$

$$\frac{\partial \varepsilon'_{i+1}}{\partial a} = \delta \exp^{-a\delta} r_{t_i} + \frac{b}{a^2} (1 - \exp^{-a\delta}) - \frac{\delta b}{a} \exp^{-a\delta}$$

$$\frac{\partial \varepsilon'_{i+1}}{\partial b} = -\frac{1}{a} (1 - \exp^{-a\delta})$$

$$\frac{\partial \varepsilon'_{i+1}}{\partial \sigma} = 0$$

Les éléments de la **Hessienne des résidus** :

$$\frac{\partial^2 \varepsilon'_{i+1}}{\partial a^2} = -\delta^2 \exp^{-a\delta} r_{t_i} - \frac{2b}{a^3} (1 - \exp^{-a\delta}) + \frac{2\delta b}{a^2} \exp^{-a\delta} + \frac{\delta^2 b}{a} \exp^{-a\delta} ;$$

$$\frac{\partial^2 \varepsilon'_{i+1}}{\partial a \partial b} = \frac{\partial^2 \varepsilon'_{i+1}}{\partial b \partial a} = \frac{(1 - \exp^{-a\delta})}{a^2} - \frac{\delta}{a} \exp^{-a\delta} ;$$

$$\frac{\partial^2 \varepsilon'_{i+1}}{\partial a \partial \sigma} = \frac{\partial^2 \varepsilon'_{i+1}}{\partial \sigma \partial a} = 0 ;$$

$$\frac{\partial^2 \varepsilon'_{i+1}}{\partial b^2} = 0 ;$$

$$\frac{\partial^2 \varepsilon'_{i+1}}{\partial b \partial \sigma} = \frac{\partial^2 \varepsilon'_{i+1}}{\partial \sigma \partial b} = 0 ;$$

$$\frac{\partial^2 \varepsilon'_{i+1}}{\partial \sigma^2} = 0$$

Dérivées de la fonction log vraisemblance :

$$\frac{\partial \Lambda_{i+1}}{\partial a} = -\frac{\delta \exp^{-2a\delta}}{1 - \exp^{-2a\delta}} + \frac{0.5}{a} - \frac{\varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} + \frac{2a \delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} - \frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})}$$

$$\frac{\partial \Lambda_{i+1}}{\partial b} = -\frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial b} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})}$$

$$\frac{\partial \Lambda_{i+1}}{\partial \sigma} = -\frac{1}{\sigma} + \frac{2a \varepsilon'^2_{i+1}}{\sigma^3 (1 - \exp^{-2a\delta})}$$

Les éléments de la **Hessienne pour la log vraisemblance** :

Calculs intermédiaires :

$$\left(2a \delta \exp^{-2a\delta} \varepsilon'^2_{i+1} \right)' = \left(2\delta \exp^{-2a\delta} - 4a\delta^2 \exp^{-2a\delta} \right) \varepsilon'^2_{i+1} + 4a\delta \exp^{-2a\delta} \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}$$

$$\left(\sigma^2 (1 - \exp^{-2a\delta})^2 \right)' = \sigma^2 \left(4\delta \exp^{-2a\delta} (1 - \exp^{-2a\delta}) \right)$$

$$\left(\frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} \right)' = \frac{\left(\left(2 \frac{\partial \varepsilon'_{i+1}}{\partial a} + 2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial a^2} \right) \varepsilon'_{i+1} + 2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \frac{\partial \varepsilon'_{i+1}}{\partial a} \right) \sigma^2 (1 - \exp^{-2a\delta}) - 2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1} * 2\delta \sigma^2 \exp^{-2a\delta}}{(\sigma^2 (1 - \exp^{-2a\delta}))^2}$$

$$\begin{aligned} \frac{\partial^2 \Lambda_{i+1}}{\partial a^2} = & - \left(\frac{-2\delta^2 \exp^{-2a\delta} (1 - \exp^{-2a\delta}) - \delta \exp^{-2a\delta} (2\delta \exp^{-2a\delta})}{(1 - \exp^{-2a\delta})^2} \right) - \frac{0.5}{a^2} \\ & - \left(\frac{2 \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1} \sigma^2 (1 - \exp^{-2a\delta}) - \varepsilon'^2_{i+1} \sigma^2 (2\delta \exp^{-2a\delta})}{(\sigma^2 (1 - \exp^{-2a\delta}))^2} \right) \\ & + \left(\frac{\left[(2\delta \exp^{-2a\delta} - 4a\delta^2 \exp^{-2a\delta}) \varepsilon'^2_{i+1} + 4a\delta \exp^{-2a\delta} \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1} \right] \sigma^2 (1 - \exp^{-2a\delta})^2 - 2a\delta \exp^{-2a\delta} \varepsilon'^2_{i+1} \sigma^2 (4\delta \exp^{-2a\delta} (1 - \exp^{-2a\delta}))}{(\sigma^2 (1 - \exp^{-2a\delta}))^2} \right) \\ & - \left(\frac{\left(\left(2 \frac{\partial \varepsilon'_{i+1}}{\partial a} + 2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial a^2} \right) \varepsilon'_{i+1} + 2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \frac{\partial \varepsilon'_{i+1}}{\partial a} \right) \sigma^2 (1 - \exp^{-2a\delta}) - 2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1} * 2\delta \sigma^2 \exp^{-2a\delta}}{(\sigma^2 (1 - \exp^{-2a\delta}))^2} \right) \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \Lambda_{i+1}}{\partial a^2} = & \frac{2\delta^2 \exp^{-2a\delta}}{(1 - \exp^{-2a\delta})} + \frac{2\delta^2 \exp^{-4a\delta}}{(1 - \exp^{-2a\delta})^2} - \frac{0.5}{a^2} - \frac{2 \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} + \frac{2\delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} \\ & + \frac{2\delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} - \frac{4a\delta^2 \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} + \frac{4a\delta \exp^{-2a\delta} \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} - \frac{2a\delta \exp^{-2a\delta} \varepsilon'^2_{i+1} \sigma^2 (4\delta \exp^{-2a\delta} (1 - \exp^{-2a\delta}))}{(\sigma^2 (1 - \exp^{-2a\delta}))^2} \\ & - \frac{\left(\left(2 \frac{\partial \varepsilon'_{i+1}}{\partial a} + 2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial a^2} \right) \varepsilon'_{i+1} + 2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \frac{\partial \varepsilon'_{i+1}}{\partial a} \right) \sigma^2 (1 - \exp^{-2a\delta})}{\sigma^2 (1 - \exp^{-2a\delta})} + \frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1} * 2\delta \sigma^2 \exp^{-2a\delta}}{(\sigma^2 (1 - \exp^{-2a\delta}))^2} \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \Lambda_{i+1}}{\partial a^2} = & \frac{2\delta^2 \exp^{-2a\delta}}{(1 - \exp^{-2a\delta})} + \frac{2\delta^2 \exp^{-4a\delta}}{(1 - \exp^{-2a\delta})^2} - \frac{0.5}{a^2} - \frac{2 \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} + \frac{2\delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} \\ & + \frac{2\delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} - \frac{4a\delta^2 \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} + \frac{4a\delta \exp^{-2a\delta} \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} - \frac{8a\delta^2 \exp^{-4a\delta} \varepsilon'^2_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^3} \\ & - \frac{2 \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} - \frac{2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial a^2} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} - \frac{2a \left(\frac{\partial \varepsilon'_{i+1}}{\partial a} \right)^2}{\sigma^2 (1 - \exp^{-2a\delta})} + \frac{4a\delta \exp^{-2a\delta} \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})^2} \end{aligned}$$

$$\frac{\partial \Lambda_{i+1}}{\partial a} = -\frac{\delta \exp^{-2a\delta}}{1 - \exp^{-2a\delta}} + \frac{0.5}{a} - \frac{\varepsilon'^2_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})} + \frac{2a\delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})^2} - \frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})}$$

$$\frac{\partial^2 \Lambda_{i+1}}{\partial a \partial b} = \frac{\partial^2 \Lambda_{i+1}}{\partial b \partial a} = -\frac{2 \frac{\partial \varepsilon'_{i+1}}{\partial b} \varepsilon'_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})} + \frac{4a\delta \exp^{-2a\delta} \frac{\partial \varepsilon'_{i+1}}{\partial b} \varepsilon'_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})^2} - \frac{2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial a \partial b} \varepsilon'_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})} - \frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \frac{\partial \varepsilon'_{i+1}}{\partial b}}{\sigma^2(1 - \exp^{-2a\delta})}$$

$$\begin{aligned} \frac{\partial^2 \Lambda_{i+1}}{\partial a \partial \sigma} = & -\frac{2 \frac{\partial \varepsilon'_{i+1}}{\partial \sigma} \varepsilon'_{i+1} \sigma^2(1 - \exp^{-2a\delta}) - 2\varepsilon'^2_{i+1} \sigma(1 - \exp^{-2a\delta})}{(\sigma^2(1 - \exp^{-2a\delta}))^2} \\ & + \frac{4a\delta \exp^{-2a\delta} \frac{\partial \varepsilon'_{i+1}}{\partial \sigma} \varepsilon'_{i+1} \sigma^2(1 - \exp^{-2a\delta})^2 - 4a\delta \exp^{-2a\delta} \varepsilon'^2_{i+1} \sigma(1 - \exp^{-2a\delta})^2}{(\sigma^2(1 - \exp^{-2a\delta}))^2} \\ & - \frac{\left(2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial a \partial \sigma} \varepsilon'_{i+1} + 2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \frac{\partial \varepsilon'_{i+1}}{\partial \sigma}\right) \sigma^2(1 - \exp^{-2a\delta}) - 4a \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1} \sigma(1 - \exp^{-2a\delta})}{(\sigma^2(1 - \exp^{-2a\delta}))^2} \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \Lambda_{i+1}}{\partial a \partial \sigma} = & -\frac{2 \frac{\partial \varepsilon'_{i+1}}{\partial \sigma} \varepsilon'_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})} + \frac{2\varepsilon'^2_{i+1}}{\sigma^3(1 - \exp^{-2a\delta})} + \frac{4a\delta \exp^{-2a\delta} \frac{\partial \varepsilon'_{i+1}}{\partial \sigma} \varepsilon'_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})^2} - \frac{4a\delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^3(1 - \exp^{-2a\delta})^2} \\ & - \frac{2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial a \partial \sigma} \varepsilon'_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})} - \frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial a} \frac{\partial \varepsilon'_{i+1}}{\partial \sigma}}{\sigma^2(1 - \exp^{-2a\delta})} + \frac{4a \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^3(1 - \exp^{-2a\delta})} \end{aligned}$$

Or $\frac{\partial \varepsilon'_{i+1}}{\partial \sigma} = 0$ et $\frac{\partial^2 \varepsilon'_{i+1}}{\partial a \partial \sigma} = \frac{\partial^2 \varepsilon'_{i+1}}{\partial \sigma \partial a} = 0$, d'où :

$$\frac{\partial^2 \Lambda_{i+1}}{\partial a \partial \sigma} = \frac{2\varepsilon'^2_{i+1}}{\sigma^3(1 - \exp^{-2a\delta})} - \frac{4a\delta \exp^{-2a\delta} \varepsilon'^2_{i+1}}{\sigma^3(1 - \exp^{-2a\delta})^2} + \frac{4a \frac{\partial \varepsilon'_{i+1}}{\partial a} \varepsilon'_{i+1}}{\sigma^3(1 - \exp^{-2a\delta})}$$

$$\frac{\partial^2 \Lambda_{i+1}}{\partial b^2} = -\frac{\left(2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial b^2} \varepsilon'_{i+1} + 2a \frac{\partial \varepsilon'_{i+1}}{\partial b} \frac{\partial \varepsilon'_{i+1}}{\partial b}\right) \sigma^2(1 - \exp^{-2a\delta})}{(\sigma^2(1 - \exp^{-2a\delta}))^2}$$

$$\frac{\partial^2 \Lambda_{i+1}}{\partial b^2} = -\frac{2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial b^2} \varepsilon'_{i+1}}{\sigma^2(1 - \exp^{-2a\delta})} - \frac{2a \left(\frac{\partial \varepsilon'_{i+1}}{\partial b}\right)^2}{\sigma^2(1 - \exp^{-2a\delta})}$$

$$\frac{\partial^2 \Lambda_{i+1}}{\partial b \partial \sigma} = - \frac{\left(2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial b \partial \sigma} \varepsilon'_{i+1} + 2a \frac{\partial \varepsilon'_{i+1}}{\partial b} \frac{\partial \varepsilon'_{i+1}}{\partial \sigma} \right) \sigma^2 (1 - \exp^{-2a\delta}) - 4a \frac{\partial \varepsilon'_{i+1}}{\partial b} \varepsilon'_{i+1} \sigma (1 - \exp^{-2a\delta})}{\left(\sigma^2 (1 - \exp^{-2a\delta}) \right)^2}$$

$$\frac{\partial^2 \Lambda_{i+1}}{\partial b \partial \sigma} = - \frac{2a \frac{\partial^2 \varepsilon'_{i+1}}{\partial b \partial \sigma} \varepsilon'_{i+1}}{\sigma^2 (1 - \exp^{-2a\delta})} - \frac{2a \frac{\partial \varepsilon'_{i+1}}{\partial b} \frac{\partial \varepsilon'_{i+1}}{\partial \sigma}}{\sigma^2 (1 - \exp^{-2a\delta})} + \frac{4a \frac{\partial \varepsilon'_{i+1}}{\partial b} \varepsilon'_{i+1}}{\sigma^3 (1 - \exp^{-2a\delta})}$$

Or $\frac{\partial \varepsilon'_{i+1}}{\partial \sigma} = 0$ et $\frac{\partial^2 \varepsilon'_{i+1}}{\partial b \partial \sigma} = \frac{\partial^2 \varepsilon'_{i+1}}{\partial \sigma \partial b} = 0$, d'où :

$$\boxed{\frac{\partial^2 \Lambda_{i+1}}{\partial b \partial \sigma} = \frac{\partial^2 \Lambda_{i+1}}{\partial \sigma \partial b} = \frac{4a \frac{\partial \varepsilon'_{i+1}}{\partial b} \varepsilon'_{i+1}}{\sigma^3 (1 - \exp^{-2a\delta})}}$$

$$\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma^2} = \frac{1}{\sigma^2} + \frac{4a \frac{\partial \varepsilon'_{i+1}}{\partial \sigma} \varepsilon'_{i+1} \sigma^3 (1 - \exp^{-2a\delta}) - 6a \varepsilon'^2_{i+1} \sigma^2 (1 - \exp^{-2a\delta})}{\left(\sigma^3 (1 - \exp^{-2a\delta}) \right)^2}$$

Or $\frac{\partial \varepsilon'_{i+1}}{\partial \sigma} = 0$, d'où :

$$\boxed{\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma^2} = \frac{1}{\sigma^2} - \frac{6a \varepsilon'^2_{i+1}}{\sigma^4 (1 - \exp^{-2a\delta})}}$$

La matrice Hessienne est alors :

Il ne faut pas oublier que $\Lambda = \sum_{i=0}^{n-1} \Lambda_{i+1}$, la matrice Hessienne est :

$$H_{\Lambda}(\theta) = \begin{pmatrix} \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial a^2} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial a \partial b} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial a \partial \sigma} \right) \\ \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial b \partial a} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial b^2} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial b \partial \sigma} \right) \\ \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma \partial a} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma \partial b} \right) & \sum_{i=0}^{n-1} \left(\frac{\partial^2 \Lambda_{i+1}}{\partial \sigma^2} \right) \end{pmatrix}$$

Les matrices de variance-covariance, corrélation et les écarts-types se déduisent des formules :

$$\text{Cov}(\theta) = -H(\theta)^{-1}$$

$$\text{Corr}(X_i, X_j) = \frac{\text{Cov}(X_i, X_j)}{\sqrt{\text{Var}(X_i)} \sqrt{\text{Var}(X_j)}}$$

c) Estimation des paramètres de Vasicek à un facteur, moindre carré non linéaire

L'estimation de ces paramètres se fait de la même façon que pour le modèle Vasicek à un facteur, on utilise ici la méthode des moindres carrés non linéaires.

Pour estimer le vecteur des paramètres θ_l , on minimise :

$$\text{Min}_{\theta} \text{SCE}(\theta_l) = \text{Min}_{\theta} \sum_{i=1}^T (l_{t_{i+1}} - f(\theta_l, l_{t_i}))^2 = l(\theta_l)'l(\theta_l)$$

avec $f(\theta_l, l_{t_i}) = l_{t_i} \exp(-a_l \delta) + \mu_l [1 - \exp(-a_l \delta)]$

$$\theta_l = \begin{bmatrix} a_l \\ \mu_l \end{bmatrix} \text{ la valeur estimée des paramètres ;}$$

$l(\theta_l)$ la fonction des résidus.

La fonction SCE est minimisée en calculant le gradient, puis en l'annulant.

$$\nabla \text{SCE}(\theta_l) = \begin{bmatrix} \frac{\partial \text{SCE}(\theta_l)}{\partial a_l} \\ \frac{\partial \text{SCE}(\theta_l)}{\partial \mu_l} \end{bmatrix} = \begin{bmatrix} -2 \sum_{i=0}^{n-1} (l_{t_{i+1}} - f(\theta_l, l_{t_i})) \frac{\partial f(\theta_l, l_{t_i})}{\partial a_l} \\ -2 \sum_{i=0}^{n-1} (l_{t_{i+1}} - f(\theta_l, l_{t_i})) \frac{\partial f(\theta_l, l_{t_i})}{\partial \mu_l} \end{bmatrix}$$

Les dérivées premières de la fonction $f(\theta_l, l_{t_i})$:

$$\frac{\partial f(\theta_l, l_{t_i})}{\partial a_l} = -l_{t_i} \delta \exp(-a_l \delta) + \mu_l \delta \exp(-a_l \delta)$$

$$\frac{\partial f(\theta_l, l_{t_i})}{\partial \mu_l} = 1 - \exp(-a_l \delta)$$

Pour obtenir une solution exacte, il est nécessaire d'appliquer un changement de variables :

$$\alpha_l = \exp(-a_l \delta), \quad \beta_l = \mu_l [1 - \exp(-a_l \delta)].$$

On obtient dès lors : $f(\theta_l, l_{t_i}) = l_{t_i} \alpha_l + \beta_l$ où $\theta_l = \begin{bmatrix} \alpha_l \\ \beta_l \end{bmatrix}$

$$\begin{aligned} \frac{\partial f(\theta_l, l_{t_i})}{\partial \alpha_l} &= l_{t_i} \\ \frac{\partial f(\theta_l, l_{t_i})}{\partial \beta_l} &= 1 \end{aligned}$$

$$\nabla \text{SCE}(\theta_l) = \begin{bmatrix} -2 \sum_{i=0}^{n-1} (l_{t_{i+1}} - f(\theta_l, l_{t_i})) \frac{\partial f(\theta_l, l_{t_i})}{\partial \alpha_l} \\ -2 \sum_{i=0}^{n-1} (l_{t_{i+1}} - f(\theta_l, l_{t_i})) \frac{\partial f(\theta_l, l_{t_i})}{\partial \beta_l} \end{bmatrix} = \begin{bmatrix} -2 \sum_{i=0}^{n-1} (l_{t_{i+1}} - l_{t_i} \alpha_l - \beta_l) l_{t_i} \\ -2 \sum_{i=0}^{n-1} (l_{t_{i+1}} - l_{t_i} \alpha_l - \beta_l) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -2 \sum_{i=0}^{n-1} (l_{t_{i+1}} - f(\theta_l, l_{t_i})) \frac{\partial f(\theta_l, l_{t_i})}{\partial \alpha_l} \\ -2 \sum_{i=0}^{n-1} (l_{t_{i+1}} - f(\theta_l, l_{t_i})) \frac{\partial f(\theta_l, l_{t_i})}{\partial \beta_l} \end{bmatrix} = \begin{bmatrix} -2 \sum_{i=0}^{n-1} (l_{t_{i+1}} - l_{t_i} \alpha_l - \beta_l) l_{t_i} \\ -2 \sum_{i=0}^{n-1} (l_{t_{i+1}} - l_{t_i} \alpha_l - \beta_l) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

De la seconde équation, on déduit l'expression de β_l :

$$-2 \sum_{i=0}^{n-1} (l_{t_{i+1}} - l_{t_i} \alpha_l - \beta_l) = 0 \Rightarrow \sum_{i=0}^{n-1} (l_{t_{i+1}}) - \alpha_l \sum_{i=0}^{n-1} (l_{t_i}) - n \beta_l = 0$$

$$\boxed{\beta_l = \frac{\sum_{i=0}^{n-1} (l_{t_{i+1}}) - \alpha_l \sum_{i=0}^{n-1} (l_{t_i})}{n}}$$

Puis l'expression de α_l est déduite de la première équation et de β_l :

$$-2 \sum_{i=0}^{n-1} (l_{t_{i+1}} - l_{t_i} \alpha_l - \beta_l) l_{t_i} = 0 \Rightarrow \sum_{i=0}^{n-1} \left(l_{t_{i+1}} - l_{t_i} \alpha_l - \frac{\sum_{i=0}^{n-1} (l_{t_{i+1}}) - \alpha_l \sum_{i=0}^{n-1} (l_{t_i})}{n} \right) l_{t_i} = 0$$

$$\sum_{i=0}^{n-1} \left(l_{t_{i+1}} - l_{t_i} \alpha_l - \frac{\sum_{i=0}^{n-1} (l_{t_{i+1}}) - \alpha_l \sum_{i=0}^{n-1} (l_{t_i})}{n} \right) l_{t_i} = \sum_{i=0}^{n-1} (l_{t_{i+1}} l_{t_i}) - \alpha_l \sum_{i=0}^{n-1} (l_{t_i}^2) - \frac{\sum_{i=0}^{n-1} (l_{t_{i+1}}) \sum_{i=0}^{n-1} (l_{t_i})}{n} + \frac{\alpha_l \sum_{i=0}^{n-1} (l_{t_i})^2}{n}$$

$$\alpha_l \sum_{i=0}^{n-1} (l_{t_i}^2) - \alpha_l \frac{\sum_{i=0}^{n-1} (l_{t_i})^2}{n} = \sum_{i=0}^{n-1} (l_{t_{i+1}} l_{t_i}) - \frac{\sum_{i=0}^{n-1} (l_{t_{i+1}}) \sum_{i=0}^{n-1} (l_{t_i})}{n}$$

$$\boxed{\alpha_l = \frac{n \sum_{i=0}^{n-1} (l_{t_{i+1}} l_{t_i}) - \sum_{i=0}^{n-1} (l_{t_{i+1}}) \sum_{i=0}^{n-1} (l_{t_i})}{n \sum_{i=0}^{n-1} (l_{t_i}^2) - \sum_{i=0}^{n-1} (l_{t_i})^2}}$$

A partir de $\hat{\alpha}$ et $\hat{\beta}$, on obtient la valeur estimée des paramètres $\hat{a}$, $\hat{\mu}$:

- $\alpha = \exp^{-a\delta} \Rightarrow \hat{a} = -\frac{\ln(\hat{\alpha})}{\delta}$
- $\hat{\mu} = \frac{b}{a}(1 - \alpha) \Rightarrow \hat{\mu} = \frac{\hat{a}\hat{\beta}}{1 - \hat{\alpha}}$

Enfin, v_l :

$$v_l \varepsilon = \sigma_l \sqrt{\frac{1 - \exp^{-2a_l \delta}}{2a_l}} \Rightarrow \hat{v}_l \hat{\varepsilon} = \sigma_l \sqrt{\frac{1 - \exp^{-2\hat{a}_l \delta}}{2\hat{a}_l}}$$

$$\Rightarrow \hat{\sigma}_l = \sqrt{\frac{2\hat{a}_l}{1 - \exp^{-2\hat{a}_l\delta}} * \frac{\sum_i (\hat{l}_{i+1})^2}{n}}$$

2. Modèle de Vasicek à deux facteurs

a) Estimation des paramètres du taux court, Vasicek à deux facteurs

Pour estimer les paramètres de la dynamique du taux court terme, il faut tenir compte de l'évaluation du taux de long terme. La discrétisation du taux court terme est alors :

$$r_{t_{i+1}} = \exp(-a_r \delta) r_{t_i} + \hat{l}_{t_i} [1 - \exp(-a_r \delta)] + \sigma_r \sqrt{\frac{1 - \exp[-2a_r \delta]}{2a_r}} \varepsilon_{r,t_i}$$

Avec $\varepsilon_{r,t_i} \sim N(0,1)$.

Remarque : dans cette dynamique, il y a un paramètre de moins à estimer.

Puis on procède de la même façon que le taux long. Pour estimer le vecteur des paramètres θ_r , on minimise :

$$\text{Min}_{\theta} \text{SCE}(\theta_r) = \text{Min}_{\theta} \sum_{t=1}^T (r_{t_{i+1}} - f(\theta_r, r_{t_i}))^2 = r(\theta_r)' r(\theta_r)$$

avec $f(\theta_r, r_{t_i}) = r_{t_i} \exp(-a_r \delta) + \hat{l}_{t_i} [1 - \exp(-a_r \delta)]$

$\theta_r = \begin{bmatrix} a_r \\ \sigma_r \end{bmatrix}$ la valeur estimée des paramètres ;

$r(\theta_r)$ la fonction des résidus.

Le gradient de $f(\theta_r, r_{t_i})$:

$$\frac{\partial f(\theta_r, r_{t_i})}{\partial a_r} = -r_{t_i} \delta \exp(-a_r \delta) + \hat{l}_{t_i} \delta \exp(-a_r \delta)$$

$$\begin{aligned} \nabla \text{SCE}(\theta_r) &= -2 \sum_{i=0}^{n-1} (r_{t_{i+1}} - f(\theta_r, r_{t_i})) \frac{\partial f(\theta_r, r_{t_i})}{\partial a_r} = 0 \\ &= -2 \sum_{i=0}^{n-1} (r_{t_{i+1}} - r_{t_i} \exp(-a_r \delta) - \hat{l}_{t_i} [1 - \exp(-a_r \delta)]) (-r_{t_i} \delta \exp(-a_r \delta) + \hat{l}_{t_i} \delta \exp(-a_r \delta)) = 0 \end{aligned}$$

On applique aussi le même changement de variable, à savoir $\alpha_r = \exp(-a_r \delta)$:

$$\nabla \text{SCE}(\theta_r) = -2 \sum_{i=0}^{n-1} (r_{t_{i+1}} - r_{t_i} \alpha_r - \hat{l}_{t_i} (1 - \alpha_r)) (-r_{t_i} \delta \alpha_r + \hat{l}_{t_i} \delta \alpha_r) = 0$$

$$\begin{aligned}
& \sum_{i=0}^{n-1} (r_{t_{i+1}} - r_{t_i} \alpha_r - \hat{l}_{t_i} (1 - \alpha_r)) (-r_{t_i} \delta \alpha_r + \hat{l}_{t_i} \delta \alpha_r) \\
&= -\delta \alpha_r \sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) + \delta \alpha_r \sum_{i=0}^{n-1} (r_{t_{i+1}} \hat{l}_{t_i}) + \delta \alpha_r^2 \sum_{i=0}^{n-1} (r_{t_i}^2) - \delta \alpha_r^2 \sum_{i=0}^{n-1} (r_{t_i} \hat{l}_{t_i}) + \delta \alpha_r (1 - \alpha_r) \sum_{i=0}^{n-1} (r_{t_i} \hat{l}_{t_i}) - \delta \alpha_r (1 - \alpha_r) \sum_{i=0}^{n-1} (\hat{l}_{t_i}^2) \\
&= \delta \alpha_r^2 \left[\sum_{i=0}^{n-1} (r_{t_i}^2) - \sum_{i=0}^{n-1} (r_{t_i} \hat{l}_{t_i}) + \sum_{i=0}^{n-1} (\hat{l}_{t_i}^2) \right] - \delta \alpha_r \sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) + \delta \alpha_r \sum_{i=0}^{n-1} (r_{t_{i+1}} \hat{l}_{t_i}) + \delta \alpha_r \sum_{i=0}^{n-1} (r_{t_i} \hat{l}_{t_i}) - \delta \alpha_r \sum_{i=0}^{n-1} (\hat{l}_{t_i}^2) \\
&= \delta \alpha_r \left[\alpha_r \sum_{i=0}^{n-1} (r_{t_i} - \hat{l}_{t_i})^2 - \sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) + \sum_{i=0}^{n-1} (r_{t_{i+1}} \hat{l}_{t_i}) + \sum_{i=0}^{n-1} (r_{t_i} \hat{l}_{t_i}) - \sum_{i=0}^{n-1} (\hat{l}_{t_i}^2) \right] = 0
\end{aligned}$$

$\Rightarrow \delta \alpha_r$ est différent de zéro, l'exponentielle ne s'annulant pas.

$$\begin{aligned}
\Rightarrow \quad & \alpha_r \sum_{i=0}^{n-1} (r_{t_i} - \hat{l}_{t_i})^2 - \sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) + \sum_{i=0}^{n-1} (r_{t_{i+1}} \hat{l}_{t_i}) + \sum_{i=0}^{n-1} (r_{t_i} \hat{l}_{t_i}) - \sum_{i=0}^{n-1} (\hat{l}_{t_i}^2) = 0 \\
& \hat{\alpha}_r = \frac{\sum_{i=0}^{n-1} (r_{t_{i+1}} r_{t_i}) - \sum_{i=0}^{n-1} (r_{t_{i+1}} \hat{l}_{t_i}) - \sum_{i=0}^{n-1} (r_{t_i} \hat{l}_{t_i}) + \sum_{i=0}^{n-1} (\hat{l}_{t_i}^2)}{\sum_{i=0}^{n-1} (r_{t_i} - \hat{l}_{t_i})^2}
\end{aligned}$$

L'estimation de la volatilité ν_r :

$$\begin{aligned}
\nu_r &= \sigma_r \sqrt{\frac{1 - \exp^{-2a_r \delta}}{2a_r}} \Rightarrow \hat{\nu}_r \hat{\varepsilon} = \sigma_r \sqrt{\frac{1 - \exp^{-2\hat{a}_r \delta}}{2\hat{a}_r}} \\
&\Rightarrow \hat{\sigma}_r = \sqrt{\frac{2\hat{a}_r}{1 - \exp^{-2\hat{a}_r \delta}} * \frac{\sum_{i=1}^n (\hat{r}_{i+1})^2}{n}}
\end{aligned}$$

La matrice Hessienne est donnée par :

$$H(\theta_r) = \left[\frac{\delta^2 SCE(\theta_r)}{\delta \alpha_r^2} \right] = -2 \sum_{i=0}^{n-1} -\frac{\partial f(a_r, r_{t_i})}{\partial \alpha_r} \frac{\partial f(a_r, r_{t_i})}{\partial \alpha_r} + (r_{t_{i+1}} - f(a_r, r_{t_i})) \frac{\delta^2 f(a_r, r_{t_i})}{\delta \alpha_r^2}$$

$$\text{Avec } \frac{\delta^2 f(a_r, r_{t_i})}{\delta \alpha_r^2} = r_{t_i} \delta^2 \exp(-a_r \delta) - \hat{l}_{t_i} \delta^2 \exp(-a_r \delta)$$

3. Modèle de Merton

a) Densité du modèle de Merton

Preuve :

La réalisation d'un saut est conditionnée par les sauts ayant pu se produire par le passé. La densité de la rentabilité s'écrit :

$$f_r(x) = \sum_{n=0}^{\infty} (f_{r|n}(x) * \Pr[N_s = n])$$

Les rendements sont définis par l'expression :

$$r_{t_i} = \left(\left(\mu - \frac{\sigma^2}{2} \right) \delta + \sigma W(\delta) + \sum_{k=N(t_{i-1})}^{N(t_i)} (U(k)) \right)$$

En décomposant les rendements, on constate que le premier terme de l'expression a une moyenne de $\left(\mu - \frac{\sigma^2}{2}\right)\delta$ et une variance nulle.

Le second terme $\sigma(W(\delta))$ a une moyenne nulle et une variance de $\delta\sigma^2$ car le mouvement brownien suit une loi normale $N(0,1)$.

Enfin le dernier terme possède également une moyenne nulle car U suit une loi gaussienne centrée et de variance $n\sigma_u^2$.

Propriété

Soit X_1 et X_2 deux variables indépendantes suivant des lois normales de moyenne m_1 et m_2 , et variance σ_1^2 et σ_2^2 respectivement, alors :

X_1+X_2 suit une loi normale d'espérance m_1+m_2 et de variance $\sigma_1^2 + \sigma_2^2$.

D'où :

$$r \approx N\left(\left(\mu - \frac{\sigma^2}{2}\right)\delta; \delta\sigma^2 + n\sigma_u^2\right)$$

Les rendements suivent une loi gaussienne, il est alors possible d'en déduire la densité conditionnelle :

$$f_{r|n}(x) = \frac{1}{\sqrt{2\pi * (\delta\sigma^2 + n\sigma_u^2)}} \exp\left(-\frac{1}{2 * (\delta\sigma^2 + n\sigma_u^2)} * \left(x - \left(\mu - \frac{\sigma^2}{2}\right)\delta\right)^2\right)$$

La probabilité d'avoir n réalisations de sauts : $\Pr[N_t = n] = e^{-\lambda} \frac{\lambda^n}{n!}$ dans une année, donc

$\Pr[N_t = n] = e^{-\delta\lambda} \frac{(\lambda\delta)^n}{n!}$ pour un pas de temps δ .

$$D'où : f_r(x) = \frac{e^{-\lambda\delta}}{\sqrt{2\pi}} \left[\sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n! * \sqrt{(\sigma^2\delta + n\sigma_u^2)}} * \exp\left(-\frac{\left(x - \left(\mu - \frac{\sigma^2}{2}\right)\delta\right)^2}{2 * (\sigma^2\delta + n\sigma_u^2)}\right) \right) \right]$$

b) Expression des moments

Preuve :

$$\begin{aligned} m_{2k} &= E[(r - m)^{2k}] = \int_{-\infty}^{+\infty} (x - m)^{2k} f_r(x) dx \\ &= \int_{-\infty}^{+\infty} (x - m)^{2k} \frac{e^{-\lambda\delta}}{\sqrt{2\pi}} \left[\sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n! \sigma_n} * \exp\left(-\frac{(x - m)^2}{2\sigma_n^2}\right) \right) \right] dx \\ &= \sum_{n=0}^{\infty} \left(\int_{-\infty}^{+\infty} (x - m)^{2k} \frac{\lambda^n e^{-\lambda}}{n!} * \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left(-\frac{(x - m)^2}{2 * \sigma_n^2}\right) dx \right) \end{aligned}$$

On pose : $u = \frac{1}{\sqrt{(\sigma^2 \delta + n \sigma_u^2)}} (x - m)$, soit $(x - m) = u \sqrt{(\sigma^2 \delta + n \sigma_u^2)}$

$$du = \frac{1}{\sqrt{(\sigma^2 \delta + n \sigma_u^2)}} dx$$

$$\begin{aligned} m_{2k} &= E[u^{2k}] = \int_{-\infty}^{+\infty} u^{2k} f_r(u) du \\ &= \int_{-\infty}^{+\infty} (u)^{2k} \frac{e^{-\lambda \delta}}{\sqrt{2\pi}} \left[\sum_{n=0}^{\infty} \left(\frac{(\lambda \delta)^n}{n!} (\sigma^2 \delta + n \sigma_u^2)^k \exp\left(-\frac{u^2}{2}\right) \right) \right] du \\ &= \sum_{n=0}^{\infty} \left(\int_{-\infty}^{+\infty} (u)^{2k} \frac{(\lambda \delta)^n e^{-\lambda \delta}}{n!} * \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) (\sigma^2 \delta + n \sigma_u^2)^k du \right) \end{aligned}$$

Par intégration par partie, on a $\int_{-\infty}^{+\infty} \frac{(u)^{2k}}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du = \frac{2k!}{2^k k!}$

$$\begin{aligned} m_{2k} &= \sum_{n=0}^{\infty} \left(\frac{(\lambda \delta)^n e^{-\lambda \delta}}{n!} (\sigma^2 \delta + n \sigma_u^2)^k \int_{-\infty}^{+\infty} \frac{(u)^{2k}}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du \right) \\ m_{2k} &= \frac{2k!}{2^k k!} \sum_{n=0}^{\infty} \frac{(\delta \lambda)^n e^{(-\lambda \delta)}}{n!} (\delta \sigma^2 + n \sigma_u^2)^k \end{aligned}$$

Calculons les moments théoriques d'ordre 2, 4 et 6 :

Rappel mathématique : $\sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^\lambda$ si $|\lambda| < 1$

Cas k=1

$$\begin{aligned} m_2 &= Var(r) = \frac{2!}{2^1 1!} \sum_{n=0}^{\infty} \frac{(\delta \lambda)^n e^{(-\lambda \delta)}}{n!} (\delta \sigma^2 + n \sigma_u^2)^2 \\ &= e^{(-\lambda \delta)} \left(\delta \sigma^2 \sum_{n=0}^{\infty} \frac{(\delta \lambda)^n}{n!} + \lambda \delta \sigma_u^2 \sum_{n=1}^{\infty} \frac{(\delta \lambda)^{n-1}}{(n-1)!} \right) \\ &= e^{(-\lambda \delta)} (\delta \sigma^2 e^{(\lambda \delta)} + \lambda \delta \sigma_u^2 e^{(\lambda \delta)}) \\ &= \delta (\sigma^2 + \lambda \sigma_u^2) \end{aligned}$$

Cas k=2

$$\begin{aligned}
 m_4 &= \frac{4!}{2^2 2!} \sum_{n=0}^{\infty} \frac{(\delta\lambda)^n e^{(-\lambda\delta)}}{n!} (\delta\sigma^2 + n\sigma_u^2)^2 \\
 &= 3e^{(-\lambda\delta)} \left(\sum_{n=0}^{\infty} \frac{(\delta\lambda)^n}{n!} (\delta^2\sigma^4 + 2n\delta\sigma^2\sigma_u^2 + n^2\sigma_u^4) \right) \\
 &= 3e^{(-\lambda\delta)} \left(\sum_{n=0}^{\infty} \frac{(\delta\lambda)^n}{n!} (\delta^2\sigma^4 + 2n\delta\sigma^2\sigma_u^2 + n(n-1)\sigma_u^4 + n\sigma_u^4) \right) \\
 &= 3e^{(-\lambda\delta)} (\delta^2\sigma^4 e^{\lambda\delta} + 2\lambda\delta^2\sigma^2\sigma_u^2 e^{\lambda\delta} + \lambda^2\delta^2\sigma_u^4 e^{\lambda\delta} + \lambda\delta\sigma_u^4 e^{\lambda\delta}) \\
 &= 3\delta^2\sigma^4 + 6\lambda\delta^2\sigma^2\sigma_u^2 + 3\lambda^2\delta^2\sigma_u^4 + 3\lambda\delta\sigma_u^4
 \end{aligned}$$

Cas k=3

$$\begin{aligned}
 m_6 &= \frac{6!}{2^3 3!} \sum_{n=0}^{\infty} \frac{(\delta\lambda)^n e^{(-\lambda\delta)}}{n!} (\delta\sigma^2 + n\sigma_u^2)^3 \\
 &= 15e^{(-\lambda\delta)} \left(\sum_{n=0}^{\infty} \frac{(\delta\lambda)^n}{n!} (\delta^3\sigma^6 + 3\delta^2\sigma^4 n\sigma_u^2 + 3\delta\sigma^2 n^2\sigma_u^4 + n^3\sigma_u^6) \right) \\
 &= 15e^{(-\lambda\delta)} \left(\sum_{n=0}^{\infty} \frac{(\delta\lambda)^n}{n!} (\delta^3\sigma^6 + 3\delta^2\sigma^4 n\sigma_u^2 + n(n-1)3\delta\sigma^2\sigma_u^4 + n3\delta\sigma^2\sigma_u^4 + n(n-1)(n-2)\sigma_u^6 + 3n(n-1)\sigma_u^6 + n\sigma_u^6) \right) \\
 &= 15e^{(-\lambda\delta)} (\delta^3\sigma^6 e^{\lambda\delta} + 3\lambda\delta^3\sigma^4\sigma_u^2 e^{\lambda\delta} + 3\lambda^2\delta^3\sigma^2\sigma_u^4 e^{\lambda\delta} + 3\lambda\delta^2\sigma^2\sigma_u^4 e^{\lambda\delta} + \lambda^3\delta^3\sigma_u^6 e^{\lambda\delta} + 3\lambda^2\delta^2\sigma_u^6 e^{\lambda\delta} + \lambda\delta\sigma_u^6 e^{\lambda\delta}) \\
 &= 15\delta^3\sigma^6 + 45\lambda\delta^3\sigma^4\sigma_u^2 + 45\lambda^2\delta^3\sigma^2\sigma_u^4 + 45\lambda\delta^2\sigma^2\sigma_u^4 + 15\lambda^3\delta^3\sigma_u^6 + 45\lambda^2\delta^2\sigma_u^6 + 15\lambda\delta\sigma_u^6
 \end{aligned}$$

c) Estimation des paramètres par le maximum de vraisemblance

L'objectif est de trouver les paramètres qui permettent de maximiser la vraisemblance d'obtenir les valeurs de l'échantillon de données.

Cette méthode s'appuie sur la définition de densité de probabilité de la variable aléatoire.

Nous disposons d'un échantillon de n données indépendantes : $X_1, X_2, \dots, X_n$.

X est une variable aléatoire de densité de probabilité $f(x, \theta)$ où θ est un vecteur de p paramètres, alors la fonction du max de vraisemblance est donnée par :

$$\begin{aligned}
 L(\theta) &= \prod_{i=1}^n f(r_i; \theta) \\
 L(\theta) &= \prod_{i=1}^n f_r(r_i, \lambda, \mu, \sigma^2, \sigma_u^2) = \prod_{i=1}^n \frac{e^{-\lambda\delta}}{\sqrt{2\pi}} \left[\sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n! \sqrt{(\delta\sigma^2 + n\sigma_u^2)}} * \exp \left(-\frac{\left(r_i - \left(\mu - \frac{\sigma^2}{2} \right) \delta \right)^2}{2 * (\delta\sigma^2 + n\sigma_u^2)} \right) \right) \right]
 \end{aligned}$$

Avec $\theta = [\lambda, \mu, \sigma^2, \sigma_u^2]$

En pratique, on maximise le logarithme de la fonction de vraisemblance :

$$\Lambda = \log(L(\theta)) = \sum_{i=1}^n \log f(r_i; \theta)$$

$$\log(L(\theta)) = \sum_{i=1}^n \frac{e^{-\lambda\delta}}{\sqrt{2\pi}} \left[\sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n! \sqrt{(\delta\sigma^2 + n\sigma_u^2)}} * \exp \left(-\frac{\left(r_i - \left(\mu - \frac{\sigma^2}{2} \right) \delta \right)^2}{2 * (\delta\sigma^2 + n\sigma_u^2)} \right) \right) \right]$$

$$\Lambda = N \left(-\lambda\delta - \log(\sqrt{2\pi}) \right) + \sum_{i=1}^N \log \left[\sum_{n=0}^{\infty} \left(\frac{(\lambda\delta)^n}{n! \sqrt{(\delta\sigma^2 + n\sigma_u^2)}} * \exp \left(-\frac{\left(r_i - \left(\mu - \frac{\sigma^2}{2} \right) \delta \right)^2}{2 * (\delta\sigma^2 + n\sigma_u^2)} \right) \right) \right]$$

Remarque : lorsque $\sigma_u^2 = 0$, il n'y a plus de saut, on se ramène au modèle de Black et Scholes.

Pour maximiser la log vraisemblance, il faut calculer les dérivées premières. Les dérivées secondes vont être utiles pour déterminer la variance des estimateurs qui est : $Cov(\theta) = -H(\theta)^{-1}$

Pour faciliter l'écriture et la lecture du calcul, posons :

$$g(r_{t_i}) = \sum_{n=0}^{\infty} \left[\frac{(\lambda\delta)^n}{n! \sqrt{\delta\sigma^2 + n\sigma_u^2}} \exp \left(-\frac{\left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)^2}{2(\delta\sigma^2 + n\sigma_u^2)} \right) \right]$$

$$v(r_{t_i}) = -\frac{\left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)^2}{2(\delta\sigma^2 + n\sigma_u^2)}$$

$$\text{D'où } g(r_{t_i}) = \sum_{n=0}^{\infty} \left[\frac{(\lambda\delta)^n}{n! \sqrt{\delta\sigma^2 + n\sigma_u^2}} \exp(v(r_{t_i})) \right]$$

$$\text{Donc } \Lambda = -N \left(\lambda\delta + \log(\sqrt{2\pi}) \right) + \sum_{i=1}^N \log(g(r_{t_i}))$$

Calcul intermédiaire des dérivées premières de v :

$$\frac{\partial v}{\partial \lambda} = 0$$

$$\frac{\partial v}{\partial \mu} = -2\delta(-1) \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right) * \frac{1}{2(\delta\sigma^2 + n\sigma_u^2)}$$

$$= \frac{\delta \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta\sigma^2 + n\sigma_u^2)}$$

$$\begin{aligned}
\frac{\partial v}{\partial \sigma} &= -\frac{2\left(\frac{1}{2} * 2\delta\sigma\right)\left(r_{t_i} - \delta\left(\mu - \frac{\sigma^2}{2}\right)\right)}{2(\delta\sigma^2 + n\sigma_u^2)} - (-1)\frac{2 * 2\delta\sigma\left(r_{t_i} - \delta\left(\mu - \frac{\sigma^2}{2}\right)\right)^2}{4(\delta\sigma^2 + n\sigma_u^2)^2} \\
&= -\frac{\delta\sigma\left(r_{t_i} - \delta\left(\mu - \frac{\sigma^2}{2}\right)\right)}{(\delta\sigma^2 + n\sigma_u^2)} + \frac{\delta\sigma\left(r_{t_i} - \delta\left(\mu - \frac{\sigma^2}{2}\right)\right)^2}{(\delta\sigma^2 + n\sigma_u^2)^2} \\
&= \frac{\delta\sigma\left(r_{t_i} - \delta\left(\mu - \frac{\sigma^2}{2}\right)\right)}{(\delta\sigma^2 + n\sigma_u^2)} \left[-1 + \frac{\left(r_{t_i} - \delta\left(\mu - \frac{\sigma^2}{2}\right)\right)}{(\delta\sigma^2 + n\sigma_u^2)} \right] \\
\frac{\partial v}{\partial \sigma_u} &= -(-1)\frac{4n\sigma_u\left(r_{t_i} - \delta\left(\mu - \frac{\sigma^2}{2}\right)\right)^2}{4(\delta\sigma^2 + n\sigma_u^2)^2} \\
&= \frac{n\sigma_u\left(r_{t_i} - \delta\left(\mu - \frac{\sigma^2}{2}\right)\right)^2}{(\delta\sigma^2 + n\sigma_u^2)^2}
\end{aligned}$$

Et les dérivées secondes de v :

$$\begin{aligned}
\frac{\partial^2 v}{\partial^2 \lambda} &= \frac{\partial^2 v}{\partial \lambda \partial \mu} = \frac{\partial^2 v}{\partial \lambda \partial \sigma} = \frac{\partial^2 v}{\partial \lambda \partial \sigma_u} = 0 \\
\frac{\partial^2 v}{\partial^2 \mu} &= \frac{-\delta}{(\delta\sigma^2 + n\sigma_u^2)} \\
\frac{\partial^2 v}{\partial \mu \partial \sigma} &= \frac{\frac{1}{2} * 2\delta^2 \sigma}{(\delta\sigma^2 + n\sigma_u^2)} - \frac{2\delta^2 \sigma\left(r_{t_i} - \delta\left(\mu - \frac{\sigma^2}{2}\right)\right)}{(\delta\sigma^2 + n\sigma_u^2)^2} \\
&= \frac{\delta^2 \sigma}{(\delta\sigma^2 + n\sigma_u^2)} \left[1 - \frac{2\left(r_{t_i} - \delta\left(\mu - \frac{\sigma^2}{2}\right)\right)}{(\delta\sigma^2 + n\sigma_u^2)} \right] \\
\frac{\partial^2 v}{\partial \mu \partial \sigma_u} &= -\frac{2n\delta\sigma_u\left(r_{t_i} - \delta\left(\mu - \frac{\sigma^2}{2}\right)\right)}{(\delta\sigma^2 + n\sigma_u^2)^2} \\
\frac{\partial^2 v}{\partial \sigma \partial \mu} &= \frac{\partial^2 v}{\partial \mu \partial \sigma}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 v}{\partial^2 \sigma} &= \left[\frac{\delta^2 \sigma^2}{(\delta \sigma^2 + n \sigma_u^2)} + \frac{\delta \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)} - \frac{2 \delta^2 \sigma^2 \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)^2} \right] \left[-1 + \frac{\delta}{(\delta \sigma^2 + n \sigma_u^2)} \right] + \\
&\frac{\delta \sigma \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)} \left[\frac{\frac{1}{2} * 2 \delta \sigma}{(\delta \sigma^2 + n \sigma_u^2)} - \frac{2 \delta \sigma \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)^2} \right] \\
&= - \frac{\delta^2 \sigma^2}{(\delta \sigma^2 + n \sigma_u^2)} - \frac{\delta \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)} + \frac{4 \delta^2 \sigma^2 \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)^2} \\
&+ \frac{\delta \sigma \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)^2}{(\delta \sigma^2 + n \sigma_u^2)^2} - \frac{4 \delta^2 \sigma^2 \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)^2}{(\delta \sigma^2 + n \sigma_u^2)^3} \\
\frac{\partial^2 v}{\partial \sigma \partial \sigma_u} &= \frac{-2 n \delta \sigma \sigma_u \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)^2} \left[-1 + \frac{\delta}{(\delta \sigma^2 + n \sigma_u^2)} \right] + \\
&\frac{\delta \sigma \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)} - \frac{2 n \sigma_u \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)^2} \\
&= \frac{2 n \delta \sigma \sigma_u \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)^2} \left[1 - \frac{\left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)} - \frac{\left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)} \right] \\
&= \frac{2 n \delta \sigma \sigma_u \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)^2} \left[1 - 2 \frac{\left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)}{(\delta \sigma^2 + n \sigma_u^2)} \right] \\
\frac{\partial^2 v}{\partial \sigma_u \partial \mu} &= \frac{\partial^2 v}{\partial \mu \partial \sigma_u} \\
\frac{\partial^2 v}{\partial \sigma_u \partial \sigma} &= \frac{\partial^2 v}{\partial \sigma \partial \sigma_u}
\end{aligned}$$

$$\begin{aligned} \frac{\partial^2 v}{\partial \sigma_u^2} &= \frac{n \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)^2}{(\delta \sigma^2 + n \sigma_u^2)^2} - \frac{4n^2 \sigma_u^2 \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)^2}{(\delta \sigma^2 + n \sigma_u^2)^3} \\ &= \frac{n \left(r_{t_i} - \delta \left(\mu - \frac{\sigma^2}{2} \right) \right)^2}{(\delta \sigma^2 + n \sigma_u^2)^2} \left[1 - \frac{4n^2 \sigma_u^2}{(\delta \sigma^2 + n \sigma_u^2)} \right] \end{aligned}$$

On en déduit les dérivées premières de $g(r_{t_i})$:

$$\begin{aligned} \frac{\partial g}{\partial \mu} &= \sum_{n=0}^{\infty} \left[\frac{(\lambda \delta)^n}{n! \sqrt{\delta \sigma^2 + n \sigma_u^2}} \frac{\partial v}{\partial \mu} \exp(v) \right] \\ \frac{\partial g}{\partial \sigma} &= \sum_{n=0}^{\infty} \left[-\frac{\sigma \delta (\lambda \delta)^n}{n! (\delta \sigma^2 + n \sigma_u^2)^{3/2}} \exp(v) \right] + \sum_{n=0}^{\infty} \left[\frac{(\lambda \delta)^n}{n! \sqrt{\delta \sigma^2 + n \sigma_u^2}} \frac{\partial v}{\partial \sigma} \exp(v) \right] \\ \frac{\partial g}{\partial \lambda} &= \sum_{n=0}^{\infty} \left[\frac{n \delta^n \lambda^{n-1}}{n! \sqrt{\delta \sigma^2 + n \sigma_u^2}} \exp(v) \right] \\ \frac{\partial g}{\partial \sigma_u} &= \sum_{n=0}^{\infty} \left[\frac{-n \sigma_u (\lambda \delta)^n}{n! (\delta \sigma^2 + n \sigma_u^2)^{3/2}} \exp(v) \right] + \sum_{n=0}^{\infty} \left[\frac{(\lambda \delta)^n}{n! \sqrt{\delta \sigma^2 + n \sigma_u^2}} \frac{\partial v}{\partial \sigma_u} \exp(v) \right] \end{aligned}$$

Et les dérivées secondes de $g(r_{t_i})$:

$$\begin{aligned} \frac{\partial^2 g}{\partial \mu^2} &= \sum_{n=0}^{\infty} \left[\frac{(\lambda \delta)^n}{n! \sqrt{\delta \sigma^2 + n \sigma_u^2}} \exp(v) \left(\frac{\partial^2 v}{\partial \mu^2} + \frac{\partial v}{\partial \mu} \frac{\partial v}{\partial \mu} \right) \right] \\ \frac{\partial g}{\partial \mu \partial \sigma} &= \sum_{n=0}^{\infty} \left[-\frac{\sigma \delta (\lambda \delta)^n}{n! (\delta \sigma^2 + n \sigma_u^2)^{3/2}} \frac{\partial v}{\partial \mu} \exp(v) + \frac{(\lambda \delta)^n}{n! \sqrt{\delta \sigma^2 + n \sigma_u^2}} \exp(v) \left(\frac{\delta^2 v}{\delta \mu \partial \sigma} + \frac{\partial v}{\partial \mu} \frac{\partial v}{\partial \sigma} \right) \right] \\ \frac{\partial^2 g}{\partial \mu \partial \lambda} &= \sum_{n=0}^{\infty} \left[\frac{n \delta^n \lambda^{n-1}}{n! \sqrt{\delta \sigma^2 + n \sigma_u^2}} \frac{\partial v}{\partial \mu} \exp(v) \right] \\ \frac{\partial g}{\partial \mu \partial \sigma_u} &= \sum_{n=0}^{\infty} \left[-\frac{n \sigma_u (\lambda \delta)^n}{n! (\delta \sigma^2 + n \sigma_u^2)^{3/2}} \frac{\partial v}{\partial \mu} \exp(v) + \frac{(\lambda \delta)^n}{n! \sqrt{\delta \sigma^2 + n \sigma_u^2}} \exp(v) \left(\frac{\delta^2 v}{\delta \mu \partial \sigma_u} + \frac{\partial v}{\partial \mu} \frac{\partial v}{\partial \sigma_u} \right) \right] \\ \frac{\partial g}{\partial \sigma \partial \mu} &= \frac{\partial g}{\partial \mu \partial \sigma} \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 g}{\partial \sigma^2} &= \sum_{n=0}^{\infty} \left[-\frac{\delta(\lambda\delta)^n}{n!(\delta\sigma^2 + n\sigma_u^2)^{3/2}} \exp(v) + \frac{3\delta^2\sigma^2(\lambda\delta)^n}{n!(\delta\sigma^2 + n\sigma_u^2)^{5/2}} \exp(v) - \frac{\sigma\delta(\lambda\delta)^n}{n!(\delta\sigma^2 + n\sigma_u^2)^{3/2}} \exp(v) \frac{\partial v}{\partial \sigma} \right] + \\
&\sum_{n=0}^{\infty} \left[\frac{\delta(\lambda\delta)^n n\sigma\delta^{n+1}}{n!(\delta\sigma^2 + n\sigma_u^2)^{3/2}} \exp(v) \frac{\partial v}{\partial \sigma} + \frac{(\lambda\delta)^n \lambda^{n-1}}{n!\sqrt{\delta\sigma^2 + n\sigma_u^2}} \exp(v) \left(\frac{\delta^2 v}{\delta\sigma^2} + \frac{\partial v}{\partial \sigma} \frac{\partial v}{\partial \sigma_u} \right) \right] \\
\frac{\partial^2 g}{\partial \sigma \partial \sigma_u} &= \sum_{n=0}^{\infty} \left[\frac{3n\sigma\delta\sigma_u(\lambda\delta)^n}{n!(\delta\sigma^2 + n\sigma_u^2)^{5/2}} \exp(v) - \frac{\sigma\delta(\lambda\delta)^n}{n!(\delta\sigma^2 + n\sigma_u^2)^{3/2}} \exp(v) \frac{\partial v}{\partial \sigma_u} \right] + \\
&\sum_{n=0}^{\infty} \left[-\frac{n\sigma_u(\lambda\delta)^n}{n!(\delta\sigma^2 + n\sigma_u^2)^{3/2}} \frac{\partial v}{\partial \sigma_u} \exp(v) + \frac{(\lambda\delta)^n}{n!\sqrt{\delta\sigma^2 + n\sigma_u^2}} \exp(v) \left(\frac{\delta^2 v}{\delta\sigma\partial\sigma_u} + \frac{\partial v}{\partial \sigma} \frac{\partial v}{\partial \sigma_u} \right) \right] \\
\frac{\partial^2 g}{\partial \lambda \partial \mu} &= \sum_{n=0}^{\infty} \left[\frac{n\delta^n \lambda^{n-1}}{n!\sqrt{\delta\sigma^2 + n\sigma_u^2}} \exp(v) \frac{\partial v}{\partial \mu} \right] \\
\frac{\partial^2 g}{\partial \lambda \partial \sigma} &= \frac{\partial^2 g}{\partial \sigma \partial \lambda} \\
\frac{\partial^2 g}{\partial \lambda^2} &= \sum_{n=0}^{\infty} \left[\frac{n(n-1)\delta^n \lambda^{n-2}}{n!\sqrt{\delta\sigma^2 + n\sigma_u^2}} \exp(v) \right] \\
\frac{\partial^2 g}{\partial \lambda \partial \sigma_u} &= \sum_{n=0}^{\infty} \left[-\frac{n^2\sigma_u\delta^n \lambda^{n-1}}{n!(\delta\sigma^2 + n\sigma_u^2)^{3/2}} \exp(v) \right] + \sum_{n=0}^{\infty} \left[\frac{n\delta^n \lambda^{n-1}}{n!\sqrt{\delta\sigma^2 + n\sigma_u^2}} \exp(v) \frac{\partial v}{\partial \sigma_u} \right] \\
\frac{\partial^2 g}{\partial \sigma_u \partial \mu} &= \frac{\partial^2 g}{\partial \mu \partial \sigma_u} \\
\frac{\partial^2 g}{\partial \sigma_u \partial \sigma} &= \frac{\partial^2 g}{\partial \sigma \partial \sigma_u} \\
\frac{\partial^2 g}{\partial \sigma_u \partial \lambda} &= \frac{\partial^2 g}{\partial \lambda \partial \sigma_u} \\
\frac{\partial^2 g}{\partial \sigma_u^2} &= \sum_{n=0}^{\infty} \left[-\frac{n(\lambda\delta)^n}{n!(\delta\sigma^2 + n\sigma_u^2)^{3/2}} \exp(v) + \frac{3n^2\sigma_u^2(\lambda\delta)^n}{n!(\delta\sigma^2 + n\sigma_u^2)^{5/2}} \exp(v) - \frac{n\sigma_u(\lambda\delta)^n}{n!(\delta\sigma^2 + n\sigma_u^2)^{3/2}} \exp(v) \frac{\partial v}{\partial \sigma_u} \right] + \\
&\sum_{n=0}^{\infty} \left[-\frac{n\sigma_u(\lambda\delta)^n}{n!(\delta\sigma^2 + n\sigma_u^2)^{3/2}} \frac{\partial v}{\partial \sigma_u} \exp(v) + \frac{(\lambda\delta)^n}{n!\sqrt{\delta\sigma^2 + n\sigma_u^2}} \exp(v) \left(\frac{\delta^2 v}{\delta\sigma_u^2} + \frac{\partial v}{\partial \sigma_u} \frac{\partial v}{\partial \sigma_u} \right) \right]
\end{aligned}$$

Enfin, on peut obtenir les dérivées premières et secondes de la log vraisemblance :

$$\frac{\partial \Lambda}{\partial \mu} = \sum_{i=1}^N \frac{\frac{\partial g}{\partial \mu}}{g} ; \quad \frac{\partial \Lambda}{\partial \sigma} = \sum_{i=1}^N \frac{\frac{\partial g}{\partial \sigma}}{g} ; \quad \frac{\partial \Lambda}{\partial \lambda} = -N\delta + \sum_{i=1}^N \frac{\frac{\partial g}{\partial \lambda}}{g} ; \quad \frac{\partial \Lambda}{\partial \sigma_u} = \sum_{i=1}^N \frac{\frac{\partial g}{\partial \sigma_u}}{g} ;$$

$$\begin{aligned}
\frac{\partial^2 \Lambda}{\partial \mu^2} &= \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \mu^2}}{g} - \frac{\frac{\partial g}{\partial \mu} \frac{\partial g}{\partial \mu}}{g^2} \right]; \quad \frac{\partial^2 \Lambda}{\partial \mu \partial \sigma} = \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \mu \partial \sigma}}{g} - \frac{\frac{\partial g}{\partial \mu} \frac{\partial g}{\partial \sigma}}{g^2} \right]; \\
\frac{\partial^2 \Lambda}{\partial \mu \partial \lambda} &= \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \mu \partial \lambda}}{g} - \frac{\frac{\partial g}{\partial \mu} \frac{\partial g}{\partial \lambda}}{g^2} \right]; \quad \frac{\partial^2 \Lambda}{\partial \mu \partial \sigma_u} = \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \mu \partial \sigma_u}}{g} - \frac{\frac{\partial g}{\partial \mu} \frac{\partial g}{\partial \sigma_u}}{g^2} \right]; \\
\frac{\partial^2 \Lambda}{\partial \sigma \partial \mu} &= \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \sigma \partial \mu}}{g} - \frac{\frac{\partial g}{\partial \sigma} \frac{\partial g}{\partial \mu}}{g^2} \right]; \quad \frac{\partial^2 \Lambda}{\partial \sigma^2} = \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \sigma^2}}{g} - \frac{\frac{\partial g}{\partial \sigma} \frac{\partial g}{\partial \sigma}}{g^2} \right]; \\
\frac{\partial^2 \Lambda}{\partial \sigma \partial \lambda} &= \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \sigma \partial \lambda}}{g} - \frac{\frac{\partial g}{\partial \sigma} \frac{\partial g}{\partial \lambda}}{g^2} \right]; \quad \frac{\partial^2 \Lambda}{\partial \sigma \partial \sigma_u} = \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \sigma \partial \sigma_u}}{g} - \frac{\frac{\partial g}{\partial \sigma} \frac{\partial g}{\partial \sigma_u}}{g^2} \right]; \\
\frac{\partial^2 \Lambda}{\partial \lambda \partial \mu} &= \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \lambda \partial \mu}}{g} - \frac{\frac{\partial g}{\partial \lambda} \frac{\partial g}{\partial \mu}}{g^2} \right]; \quad \frac{\partial^2 \Lambda}{\partial \lambda \partial \sigma} = \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \lambda \partial \sigma}}{g} - \frac{\frac{\partial g}{\partial \lambda} \frac{\partial g}{\partial \sigma}}{g^2} \right]; \\
\frac{\partial^2 \Lambda}{\partial \lambda^2} &= \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \lambda^2}}{g} - \frac{\frac{\partial g}{\partial \lambda} \frac{\partial g}{\partial \lambda}}{g^2} \right]; \quad \frac{\partial^2 \Lambda}{\partial \lambda \partial \sigma_u} = \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \lambda \partial \sigma_u}}{g} - \frac{\frac{\partial g}{\partial \lambda} \frac{\partial g}{\partial \sigma_u}}{g^2} \right]; \\
\frac{\partial^2 \Lambda}{\partial \sigma_u \partial \mu} &= \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \sigma_u \partial \mu}}{g} - \frac{\frac{\partial g}{\partial \sigma_u} \frac{\partial g}{\partial \mu}}{g^2} \right]; \quad \frac{\partial^2 \Lambda}{\partial \sigma_u \partial \sigma} = \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \sigma_u \partial \sigma}}{g} - \frac{\frac{\partial g}{\partial \sigma_u} \frac{\partial g}{\partial \sigma}}{g^2} \right]; \\
\frac{\partial^2 \Lambda}{\partial \sigma_u \partial \lambda} &= \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \sigma_u \partial \lambda}}{g} - \frac{\frac{\partial g}{\partial \sigma_u} \frac{\partial g}{\partial \lambda}}{g^2} \right]; \quad \frac{\partial^2 \Lambda}{\partial \sigma_u^2} = \sum_{i=1}^N \left[\frac{\frac{\partial^2 g}{\partial \sigma_u^2}}{g} - \frac{\frac{\partial g}{\partial \sigma_u} \frac{\partial g}{\partial \sigma_u}}{g^2} \right];
\end{aligned}$$

VIII. Références

A. Sites

- [1] Autorité de contrôle prudentiel : <http://www.banque-france.fr/acp/index.htm>
- [2] Autorité des Normes Comptables : www.anc.gouv.fr
- [3] Bloomberg : www.bloomberg.com
- [4] Compagnie Nationale des Commissaires aux Comptes : <http://www.cncc.fr/>
- [5] Direction Générale des Finances Publiques : <http://www.impots.gouv.fr/>
- [6] EIOPA : <https://eiopa.europa.eu/>
- [7] Friggit : <http://www.cgedd.developpement-durable.gouv.fr/>
- [8] IASB : <http://eifrs.iasb.org/>
- [9] IFRS : <http://www.ifrs.org>
- [10] QIS 5 : www.qis5.fr

B. Ouvrages

- [1] Brigo D., Mercurio F., Interest rate models – Theory and Practice : With smile, Inflation and Credit (2006);
- [2] Damodar N. Gujarati, Econométrie, Ouvertures économiques, (2004) ;
- [3] Davidson R., MacKinnon J.G., Estimation and Inference in Econometrics (1993);
- [4] Planchet F., Thérond P., Kamega A., Scénarios économiques en assurance, modélisation et simulation (novembre 2009) ;
- [5] Planchet F., Thérond P., Modèles financiers en Assurance (2005) ;
- [6] Planchet F., Thérond P., Jacquemin J., Modèles financiers en assurance, analyses de risqué dynamiques, (2005) ;
- [7] Portait R., Finance de marché - Instruments de base, produits dérivés, portefeuilles et risques (2008) ;

C. Documents

- [1] Ahlgrim C.k., Arcy S.P., Gorvett R.W., Modeling financial scenarios : a framework for the actuarial profession, (2005);
- [2] Ahlgrim C.k., Arcy S.P., Gorvett R.W., A comparison of actuarial financial scenario generators (2008) ;
- [3] Akaike, H., A new look at statistical model identification (1974);
- [4] Andersen, T.G., Benzoni L., Lund J., Stochastic volatility, mean drift, and jumps in the short-term interest rate, (2004) ;
- [5] Banque de France, ACP, Solvabilité 2 : principaux enseignements de la cinquième étude quantitative d'impact (QIS5) (mars 2011);
- [6] Boni E., Genot B., Momein A., Un modèle de detection des sauts dans un processus de cours boursiers (2005);
- [7] Bonnassieux M., Brunel V., Un modèle de Ho & Lee généralisé ;
- [8] Brennan M.J., Xia H., Stochastic interest rates and the bond-stock mix (2000);
- [9] Brennan M.J., Xia H., Dynamic asset allocation under inflation (2000) ;
- [10] Breusch T.S., Testing for autocorrelation in dynamic linear models, Australian economic papers, vol.17, p334-355, (1978);

- [11] Caja A., Carvas P., Rousseau V., Intégrer des informations exogènes dans un modèle d'actifs (mai 2009) ;
- [12] Cai L., Specification Testing for Multifactor Diffusion Processes: An Empirical and Methodological Analysis of Model Stability Across Different Historical Episodes (2008);
- [13] CFO Forum - European Embedded Value Principles;
- [14] Chen L., Stochastic Mean and Stochastic Volatility – A three factor Model of Term Structure of Interest Rates and its Application to the Pricing of interest Rate Derivatives (1996);
- [15] Cox J.C., Ingersoll J.E., Ross S.A., A theory of the term structure of interest rates, *Econometrica* (1985);
- [16] Durant D. (Banque de France), Guérin J. (Banque de France), Lekehal S. (ACAM), Evolutions récentes de la structure des placements des assureurs (septembre 2009);
- [17] Edgington E.S., Probability table for number of runs of signs of first differences in ordered series (1961);
- [18] Engle R., AutoRegressive Conditional Heteroscedasticity with estimates of variance of United Kingdom inflations, *Econometrica*, 50, p 987-1008 (1982);
- [19] Faleh A., Planchet F., Rulière D., Les générateurs de scénarios économiques : quelle utilisation en assurance ? (novembre 2009) ;
- [20] Godfrey L.G., Testing against general autoregressive and moving average error models when the regressor include lagged dependent variables, *Econometrica*, vol.46 (1978).
- [21] Gallant A.R., Nonlinear Regression, *American Statistician*, 29, 73-81, (1975) ;
- [22] Hardy M.R., A regime-switching model of long-term stock returns (2001) ;
- [23] Huber, A review of Wilkie's stochastic asset model (1997), B.A.I.;
- [24] Hurlin C., Tokpavi S., Evaluation des procédures de backtesting (2007) ;
- [25] Hurvich, C.M., Tsai, C.L., Regression and time series model selection in small samples (1989);
- [26] Hurvich, C.M., Tsai, C.L., Model selection for extended quasi-likelihood models in small samples (1995);
- [27] Johannes M., The statistical and Economic Role of Jumps in Continuous-time interest rate models (2004);
- [28] Jarque C.M., Bera A.K., A test for normality of observations and regression residuals (1987);
- [29] Kitts, Comments on a model of retail price inflation (1990), J.I.A.;
- [30] Kou S. G., A jump diffusion model for option pricing (2002);
- [31] Merton R.C., Option pricing when underlying stock returns are discontinuous (1976);
- [32] Planchet F., Therond P.E., L'impact de la prise en compte des sauts boursiers dans les problématiques d'assurance, (2005);
- [33] PricewaterhouseCoopers, Concept and theoretical background of Embedded Value, (16 novembre 2006);
- [34] Rambaruth G., A comparison of Wilkie-type stochastic investment models, (juillet 2003) ;
- [35] Ripley A.E., Painter I.S., Volinsky C.T., BMA : An R package for Bayesian Model Averaging (2003);
- [36] Sahin S., Cairns A., Kleinow T., Wilkie A. D., Revisiting the Wilkie investment model (juin 2008);

- [37] Schwarz, G., Estimating the dimension of a model (1978);
- [38] Thérond P., IFRS, SOLVABILITE 2, EMBEDDED VALUE : Quel traitement du risque ? ;
- [39] Vasicek O., An equilibrium characterization of the term structure (1977);
- [40] Wilkie A.D., A stochastic investment model for actuarial use (1984);
- [41] Wintenberger O., Statistique mathématique (2010);

D. Mémoires

- [1] Allag L., Modélisation et allocation stratégique d'actifs dans le cadre du référentiel de solvabilité II (2008) ;
- [2] Armel K., Structure de dépendance des générateurs de scénarios économiques, modélisation et simulation (juin 2010) ;
- [3] Bernadou J., Effet sur l'allocation d'actif du coût des fonds propres liés aux classes d'actifs dans Solvabilité II : cas des obligations convertibles (2010) ;
- [4] Kervazo V., Normes IFRS : principes et valorisation en Epargne (2009) ;
- [5] Rasqué J., Mesure et gestion du risqué de marché dans l'environnement Solvabilité II (2011) ;