



Mémoire présenté le :

pour l'obtention du Diplôme Universitaire d'actuariat de l'ISFA et l'admission à l'Institut des Actuaire

Par : Karim MZOUGH

Générateur de Scénarios Economiques : Approche par Bootstrap avec prise en compte des dépendances de queues

Confidentialité : [X] NON [ ] OUI (Durée : [ ] 1 an [ ] 2 ans)

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

Membre présents du jury de l'Institut des Actuaire

Patrick PERRET

Nordine CHOUKAR

Membres présents du jury de l'ISFA

signature

Entreprise :

Nom : PREDICA (Groupe CAA)

Signature :

Directeur de mémoire en entreprise :

Nom : Jean-Marc PHILIPPE

Signature :

Invité :

Nom :

Signature :

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

Signature du responsable entreprise

[Signature]

Signature du candidat

[Signature]

Secrétariat

Bibliothèque :

# Résumé

Depuis la réforme de Solvabilité 2, entrée en vigueur le 1er janvier 2016, le principe d'évaluation économique a conduit les entreprises d'assurance à s'intéresser particulièrement à l'évolution des grandeurs économique-financières qui conditionnent le prix des actifs financiers de leur bilan. Ces grandeurs déterminent également indirectement l'évaluation de leur passif via les interactions actif-passif principalement en assurance Vie.

La variation dans le temps des prix des actifs a donc un impact sur la richesse des entreprises d'assurance qui ont dû développer l'outil de simulation : le Générateur de Scénarios Économiques, afin de mieux appréhender leur pilotage financier et de mieux maîtriser les risques inhérents.

L'objet de ce *mémoire* est de décrire la construction d'un GSE via l'outil statistique du Bootstrap sous probabilité historique en intégrant la structure de dépendance des extrêmes des facteurs de risque.

La technique du Bootstrap, déviée de son premier usage statistique, permet de générer de nouvelles trajectoires de séries temporelles financières.

La méthodologie présentée offre l'avantage d'avoir une structure de simulation indépendante du choix économétrique de chaque actifs et du nombre d'actifs modélisés.

# Abstract

Solvency 2 reform, that came into effect on 1st January 2016, required insurers to account for their assets at their fair values. This means that insurers need to focus on economic and financial quantities that determine the price of their assets. These quantities also determine indirectly their liabilities estimates through asset-liabilities interactions mainly in life insurance.

Time variation of asset prices therefore has an impact on the wealth of insurance companies which had to develop a simulation tool - an Economic Scenarios Generator - to be in a better position to improve their financial guidance and risk management.

In this memorandum, we will describe the process to elaborate ESG using Bootstrap statistical tool under historical probability taking into account dependency structures of risk factors extremes.

We will present the Bootstrap not as a statistical tool but as a simulation tool.

Our approach is robust in the sense that we can modify the time series model and the number of considered variables without changing the whole simulation structure.

Ami, le risque est grand ; nul cas rédhibitoire ;  
Le destin est au fond de l'urne aléatoire

*Louis Honoré Fréchette 1839-1908*

# Remerciements

Après dix ans dans le domaine bancaire, j'ai souhaité m'orienter vers le monde de l'assurance en effectuant une formation continue à l'ISFA de Lyon.

Ce mémoire est l'aboutissement de ce parcours.

Je remercie mes camarades de formation continue Michel, Hong-Cat, Mustapha, Mamadou, Patrick, Cyriaque, Amilcar... pour la bonne ambiance de révision.

Je suis très reconnaissant envers le GRO<sup>1</sup> qui m'a offert la possibilité d'effectuer cette formation continue. Dans ce cadre, je remercie Patrice Bertail de m'avoir fait découvrir le bel outil du Bootstrap et ses applications. Merci à Justina Boltryck avec qui nous avons pu faire nos premières découvertes Bootstrap, je te dis dzięki<sup>2</sup>.

Je remercie l'ensemble des enseignants de l'ISFA qui m'ont fait découvrir le monde passionnant de l'Actuariat et Delphine Clemens Baines qui a géré l'organisation de cette formation continue.

Je souhaite remercier Frédéric Planchet d'avoir accepté mon sujet et de m'avoir apporté ses précieux conseils tout au long de ce mémoire. C'est un grand honneur. J'apprécie sa volonté de partager et de diffuser les connaissances et recherches en Actuariat.

Je remercie infiniment Jean-Marc Philippe, Directeur des Risques de PREDICA d'avoir accepté que je rejoigne son équipe. Je remercie les collègues pour leur soutien et en particulier Eva Limouzin qui a accepté de porter une lecture attentive à mon mémoire.

Enfin je souhaite remercier ma famille,

- mon père à qui je dédie ce mémoire et qui doit sourire là où il est.
- ma mère pour son indéfectible soutien.
- mes enfants de m'avoir prêté leur ordinateur portable Windows<sup>3</sup>.

et Stéphanie mon épouse sans qui rien n'est possible!

---

1. Groupe de Recherche Opérationnelle - Service du Crédit Agricole de prestations quantitatives, créé au Crédit Lyonnais en 1969.

2. Merci en Polonais, Nie mówię po polsku.

3. Je préférerais l'éditeur sous Windows à celui sous Mac OS pour écrire en (La)TeX.

# Synthèse

Depuis la réforme de **Solvabilité 2**, entrée en vigueur le 1er janvier 2016, le principe d'évaluation **économique** a conduit les entreprises d'assurance à s'intéresser particulièrement à l'évolution des **grandeurs économique-financières** qui conditionnent le prix des **actifs** financiers de leur bilan. Ces grandeurs déterminent également indirectement l'évaluation de leur **passif** via les interactions actif-passif principalement en assurance Vie<sup>4</sup>.

La variation dans le temps des prix des actifs a donc un impact sur la richesse des entreprises d'assurance qui ont dû développer l'outil de simulation : le **Générateur de Scénarios Économiques**, afin de mieux appréhender leur pilotage financier et de mieux maîtriser les risques inhérents.

L'objet de ce *mémoire* est de décrire la construction d'un GSE via l'outil statistique du **Bootstrap** sous probabilité **historique** en prenant en compte la **structure** de **dépendance** des **extrêmes** des facteurs de risque.

Dans la **première partie** :

1. Nous développerons les notions de probabilité **historiques** et **risque-neutres** sous lesquelles les simulations du GSE seront effectuées.  
Dans la majorité des cas les GSE utilisent les deux types de probabilité qui ne devront pas être confondus.
  - les simulations effectuées sous probabilité historique permettent de générer des trajectoires réalistes aux fins de retrouver dans les valeurs futures les propriétés statistiques observées.
  - les simulations effectuées sous probabilité risque-neutre permettent quant à elles de valoriser les flux assurantiels sur un horizon donné.
2. Nous exposerons la **nécessité** de recourir à un générateur de scénarios économiques dans les **usages** assurantiels :
  - le calcul réglementaire du ratio de solvabilité dans le cadre de la formule standard ou des approximations des formules de type simulations dans les simulations (Réplication, LSMC, ...)
  - le pilotage de la gestion actif-passif avec la prise en compte des différents risques (taux d'intérêt, crédit, ...)
  - le processus ORSA dans le cadre des calculs prospectifs.
3. Nous terminerons cette première partie par la description des **étapes universelles**/briques nécessaires à la construction d'un GSE.

---

4. En assurance Non-Vie, la durée des passifs est plus courte sauf pour certaines branches à développement long comme la responsabilité civile automobile où l'inflation devient un risque à prendre en considération pour l'évaluation des rentes.

Dans la **deuxième partie** :

Nous décrirons la technique statistique proprement dite du **Bootstrap** basée sur la distribution empirique.

Cet outil statistique a été introduit pour estimer la **variabilité** d'une statistique complexe. Il a l'avantage d'être **simple** avec une mise en place opérationnelle relativement facile. Il ne se base pas sur une hypothèse forte sur la loi sous-jacente. De plus, il a la possibilité de se généraliser aux données multi-dimensionnelles.

Dans cette partie nous insisterons sur la convergence du Bootstrap. En d'autres termes, nous montrerons que la distribution de la statistique est robuste à une petite variation par rapport à la loi sous-jacente inconnue. On pourra donc remplacer la distribution sous-jacente par la distribution empirique.

Dans la **troisième partie** :

Nous exposerons la technique du **Bootstrap** qui sera **dévié** en l'occurrence de son premier usage statistique et sera utilisé comme outil de **simulation**.

Les variables économique-financières ne présentant pas la caractéristique indispensable d'**indépendance** sous-jacente au Bootstrap, nous déterminerons les **résidus** de modèles de **séries temporelles** desdites variables qui seront **eux** l'objet d'application de la technique du Bootstrap.

Pour simuler une **nouvelle** variable, nous effectuerons donc un tirage Bootstrap des **résidus** que nous **injecterons** dans le **modèle** de série temporelle. Cette étape permettra d'obtenir de nouvelles trajectoires cohérentes modélisées par la série temporelle et basées sur le passé.

Afin que les simulations ne soient pas exclusivement basées sur l'échantillon **passé**, nous introduirons des simulations d'événements **extrêmes**.

A cet effet, nous estimerons une loi de **Pareto généralisée** (GPD) via la méthode statistique POT sur la distribution empirique des résidus. Nous **injecterons** ensuite dans le modèle de série temporelle une réalisation extrême à partir de de la loi GPD estimée pour obtenir une nouvelle variable à l'instar de ce qui est fait pour le corps de la fonction.

Dans la **quatrième partie** :

Nous avons précédemment exposé une méthodologie de simulation **univariée**. Pour pouvoir simuler plusieurs variables économique-financières nous devons introduire une **structure de dépendance** des trajectoires **multidimensionnelles**.

Dans un premier temps, l'étude portera sur les structures de dépendance habituellement utilisées dans les GSE basées sur :

- des liens souvent linéaires établis par la théorie économique
- la statistique de corrélation linéaire.

Pour notre GSE, nous choisirons une structure **implicite** et **non-paramétrique** :

- un bootstrap **vectorel** pour la partie centrale. On conserve ainsi la dépendance implicite à une date donnée entre les actifs.
- une copule **empirique** de Paul Deheuvels pour les dépendances historiques de queues.

L'avantage de cette structure est d'éviter les problématiques d'estimations paramétriques en **grande** dimension<sup>5</sup>.

Enfin, dans la **cinquième** et **dernière partie** :

Les crises **multiples** (crise des CDO, crise de la dette souveraine ...) ont affecté l'ensemble de l'environnement économique au cours de ces dernières années et introduit des comportements économique-financiers complexes à modéliser. Dans ce cadre, on constate que les hypothèses usuelles sont souvent rejetées par les tests statistiques. De ce fait, il apparaît nécessaire de revoir les modèles économétriques habituellement utilisés pour les différents actifs.

Nous explorerons quelques modèles économétriques pour les actions, le crédit, l'immobilier, les taux d'intérêt et l'inflation. A noter que pour les **taux d'intérêt**, en particulier, nous appliquerons de façon **originale** notre méthodologie de simulation, non pas directement sur la variable d'intérêt, mais sur les **paramètres** d'une fonctionnelle, celle de Nelson-Siegel.

**En conclusion** :

Notre méthodologie a l'avantage d'être **pérenne** quel que soit le choix économétrique de chaque actif et le nombre d'actifs modélisés.

La **cohérence** économique dans les trajectoires simulées des variables économique-financières est déterminée via :

- le choix économétrique pour **chaque** variable d'intérêt
- un tirage bootstrap **vectorel**
- une structure de dépendance des extrêmes basée sur les **rangs**

MOTS-CLEFS : GSE, SCR, Simulations dans les Simulations (SdS), Bootstrap, AR-ARCH, Vraisemblance, Méthode POT, Copule, Nelson-Siegel

---

5. qui correspond aux nombres d'actifs que nous souhaitons simuler.



# Executive summary

**Solvency 2** reform, that came into effect on 1st January 2016, required insurers to account for their assets at their fair values. This means that insurers need to focus on economic and financial quantities that determine the price of their assets. These quantities also determine indirectly their liabilities estimates through asset-liabilities interactions mainly in life insurance.

Time variation of asset prices therefore has an impact on the wealth of insurance companies which had to develop a simulation tool - an **Economic Scenarios Generator** - to be in a better position to improve their financial guidance and risk management.

In this memorandum, we will describe the process to elaborate ESG using Bootstrap statistical tool under **historical** probability taking into account **dependency structures** of risk factors **extremes**.

## Part 1 :

1. We will describe notions of **historical** probability and **risk-neutral** used by ESG simulations.

In most cases ESG uses both types of probabilities which should not be confused.

- simulations under historical probability enable realistic trajectories to be generated and to find stylized facts in future values.
- on the other hand, simulations under risk-neutral probability refers to market consistent pricing of insurance flows.

2. We outline the **necessity** to use an ESG in insurance **practices**

- the calculation of the Solvency Capital Requirement with the modular standard formula and/or in internal risk models (replicating portfolios, LSMC)
- asset-liabilities management taking into account different market risks (interest rate risk, credit risk, ...)
- ORSA process with prospective calculations.

3. Finally, we will describe the **usual** steps in ESG elaboration.

## Part 2 :

We will describe the concept of Bootstrap as a statistical tool based on empirical distribution.

This statistical tool was introduced to measure the standard errors of **complex** statistics. The great advantage of Bootstrap is its simplicity and can be easily implemented. Bootstrap is often used as an alternative to parametric statistical inference without knowing the underlying law. It also has the possibility to be applied with multi-dimensional data.

In this part we will insist on the consistency of the Bootstrap method.

## Part 3 :

We will present the Bootstrap not as a statistical tool but as a **simulation** tool.

As the financial quantities do not offer the i.i.d. characteristics, we will determine residuals of time series of such variables and will apply bootstrap techniques to such residuals.

To simulate a new variable we will draw a bootstrap sample of residuals and will plug it into the time series model. This step will lead to new trajectories consistent with the time series model based on the past.

To avoid a sample exclusively based on the past, we will introduce **extreme** events in simulations.

For this, we will estimate a Generalised Pareto distribution (GPD) through the POT statistical method on the empirical distribution of residuals. We will then plug it into the time series model an extreme simulation based on the estimated GPD to obtain a new variable in the same way as above.

## Part 4 :

We have just described an univariate simulation method. To simulate several financial variables we will have to introduce a **dependency structure** of multidimensional trajectories.

We will first study usual dependency structures used in ESGs based on :

- links usually appearing as linear in Economic Theory.
- correlation coefficient measure

For our ESG, we will choose an **implied** and **not** parametric structure :

- a vectorial bootstrap for the core distribution. Thus, we keep the implied dependency at a specific date between assets
- an empirical copula of Paul Deheuvels for historical dependency tails.

The advantage of using such structure is to avoid parametric inference in multi-dimensional cases.

**Part 5 :**

Recent crises (CDO crises, sovereign debt crises, ...) had an impact on the entire economical environment and introduced complex behaviors to model. In this cases the usual hypotheses are often rejected by statistical tests. It appears necessary to reconsider econometrical models usally used for different assets.

We will investigate a few alternatives time series model for equities, credit, real estate, interest rate and inflation. Please note that we will have an original approach in simulating interest rate trajectories : we will apply our methodology on the parameters of Nelson-Siegel fonctional instead of interest variables.

**Conclusion :**

The advantage of our approach is robust is the sense that we can modify the time series model and the number of considered variables without changing the whole structure.

The economic consistency of our simulated trajectories are determined by

- an independent choice of econometric for each variable
- a vectorial bootstrap draw
- a dependency structure of extremes based on ranks.

KEY-WORDS : ESG, SCR, Nested simulations (SdS), Bootstrap, AR-ARCH, Likelihood method, POT, Copula, Nelson-Siegel fonctional

# Table des matières

<b>1</b>	<b>Générateur de Scénarios Economiques</b>	<b>17</b>
1.1	Introduction . . . . .	18
1.1.1	Simulations sous probabilité historique . . . . .	20
1.1.2	Valorisation sous probabilité risque-neutre . . . . .	22
1.2	Usages du GSE en assurance . . . . .	25
1.2.1	Valorisation économique . . . . .	26
1.2.2	Calcul du SCR . . . . .	28
1.2.3	Gestion actif-passif . . . . .	40
1.2.4	ORSA . . . . .	45
1.3	Mise en oeuvre d'un GSE . . . . .	47
1.3.1	Identification des facteurs de risques . . . . .	47
1.3.2	Inputs . . . . .	48
1.3.3	Modèles de projection . . . . .	51
1.3.4	Structure de dépendance . . . . .	51
1.3.5	Probabilité de simulation . . . . .	60
1.3.6	Méthodologies d'estimation et/ou de calibration . . . . .	60
1.3.7	Validation des scénarios générés . . . . .	60
<b>2</b>	<b>Bootstrap naïf - un outil statistique</b>	<b>62</b>
2.1	Méthode d'estimation par injection : Estimateurs <i>plug-in</i> . . . . .	64
2.2	Algorithme du Bootstrap naïf d'Efron . . . . .	66
2.3	Convergence du <i>Bootstrap</i> . . . . .	67
<b>3</b>	<b>Simulations univariées</b>	<b>70</b>
3.1	Bootstrap de résidus d'un modèle économétrique . . . . .	71
3.2	Modification des queues de distributions de la loi Bootstrap . . . . .	73
3.2.1	Lois des extremums généralisés - GEV . . . . .	73
3.2.2	Méthode POT - Peaks-Over-Threshold . . . . .	75
3.2.3	Estimation de la distribution de Pareto généralisée GPD . . . . .	77
3.2.4	Loi mélange . . . . .	78
<b>4</b>	<b>Simulations multivariées</b>	<b>80</b>
4.1	Modèles intégrés . . . . .	81
4.1.1	Modèle de Wilkie - 1986 - 1995 . . . . .	81

4.1.2	Autres modèles en cascade	82
4.2	Modèles composites	83
4.2.1	Modèle d’Ahlgrim et al. 2005	84
4.2.2	Modèles à copules	85
4.2.3	Modèle avec <i>bootstrap</i> vectoriel	85
4.2.4	Simulations <i>bootstrap</i> -extrêmes	86
<b>5</b>	<b>Modélisation économétrique</b>	<b>91</b>
5.1	Modèles de séries temporelles	92
5.1.1	Bruit blanc	92
5.1.2	Processus auto-régressifs - (AR)	93
5.1.3	Processus auto-régressifs conditionnellement hétéroscédastiques - (ARCH)	95
5.1.4	Modèles avec erreurs ARCH - (ARIMA – GARCH)	97
5.1.5	Modèles à régimes	97
5.2	Tests statistiques	98
5.2.1	Tests de significativité	99
5.2.2	Tests d’adéquation	99
5.2.3	Tests d’autocorrélation sérielles des résidus	99
5.2.4	Test d’hétéroscédasticité conditionnelle	99
5.2.5	Tests de stationnarité	100
5.2.6	Critère de comparaison de modèles	100
5.3	Estimation paramétrique	101
<b>6</b>	<b>Choix économétriques pour le GSE</b>	<b>102</b>
6.1	Modèle de taux d’intérêt à paramètres dynamiques - Fonctionnelle de Nelson-Siegel	103
6.1.1	Introduction	103
6.1.2	Fonctionnelle de Nelson-Siegel	105
6.1.3	Estimation	106
6.1.4	Modélisation AR(1) des paramètres $\beta_i$	112
6.2	Modèle dynamique pour les actions	118
6.2.1	Analyse descriptive	118
6.2.2	Modélisation ARIMA – GARCH	121
6.3	Modèle dynamique pour le crédit	126
6.3.1	Analyse descriptive	126
6.3.2	Modélisation CIR	127
6.4	Modèle dynamique pour l’inflation	130
6.4.1	Analyse descriptive	130
6.4.2	Modélisation auto-régressive	131
6.4.3	Modélisation SETAR	133
6.5	Modèle dynamique pour l’immobilier	135
6.5.1	Analyse descriptive	135
6.5.2	Modélisation BS	137
6.5.3	Modélisation OU	137

<b>A</b>	<b>Copules</b>	<b>139</b>
A.1	Définition	139
A.2	Copule paramétrique elliptique	141
A.3	Copule paramétrique archimédienne	144
<b>B</b>	<b>Tests statistiques</b>	<b>149</b>
B.1	Tests de significativité	149
B.2	Tests d'adéquation	150
B.3	Tests d'auto-corrélation sérielles des résidus	152
B.4	Tests d'hétéroscédasticité conditionnelle	154
B.5	Tests de stationnarité	155
B.6	Critère de comparaison de modèles	157
<b>C</b>	<b>Estimation paramétrique</b>	<b>158</b>
C.1	Moindres carrés non-linaire	158
C.2	Quasi-Maximum de vraisemblance	160
<b>D</b>	<b>Statistiques d'ordre</b>	<b>165</b>
<b>E</b>	<b>Distribution de Pareto généralisée</b>	<b>167</b>
E.1	Graphique de l'espérance excédentaire	167
E.2	Graphique de stabilité (Hill, Pickands, ...)	169
E.3	Estimation non-paramétrique du paramètre $\xi$	173
E.4	Estimation paramétrique du paramètre de dispersion $\sigma$	173
<b>F</b>	<b>Méthodes numériques d'optimisation</b>	<b>174</b>
F.1	Algorithme à évolution différentielle	174
F.2	Algorithme de Broyden-Fletcher-Goldfarb-Shanno	175

# Notations

Ci-dessous les notations principales<sup>6</sup> de ce mémoire :

- $X := \dots$  signifie  $X$  est défini par ...
- $X \simeq \dots$  signifie est approximé par ...
- $[n]$  désigne la partie entière de  $n$ .
- $t$  la date courante de calcul ( $t = 0$  : aujourd'hui -  $t = 1$  : dans un an).
- $u$  date future ( $u > t$ )
  
- $\delta_u$  le facteur d'actualisation tel que  $\delta_u := e^{-\int_0^u r_s ds}$
- $P(t, u)$  le prix d'un zéro-coupon de maturité  $u$ .
  
- $(\mathcal{F}_t)_{t \geq 0}$  la filtration qui caractérise l'information disponible à la date  $t$  issue de trajectoires risque-neutre.
- $(\mathcal{F}_t^{RW})_{t \geq 0}$  la filtration à la date  $t$  générée par les trajectoires 33monde réel.
- $(\mathcal{F}_t^{RW-\bar{R}_i})_{t \geq 0}$  la filtration du monde réel choqué associé au risque  $R_i$  à la date  $t$
  
- $A_t$  la valeur de marché des actifs à la date  $t$ .
- $L_t$  les provisions techniques au sens SII (BE+risque de marge) à la date  $t$ .
- $\text{ANR}_t$  l'actif net réévalué à la date  $t$ .
- $\text{VIF}_t$  la *value-in-force* à la date  $t$ .
- $\text{NAV}_t$  la valeur de l'actif net (fonds propres économiques) à la date  $t$ .
  
- $F_t$  les *cash-flows* du passif (prestations, commissions, dépenses, ...) à la date  $t$ .
- $R_u$  les profits de l'entreprise d'assurance à la date  $u$ .
  
- $R_u^{R_i}$  les profits de l'entreprise d'assurance à la date  $t$  dans le scénario choqué associé au risque  $R_i$  à la date  $u$ .
  
- $\Phi(\cdot)$  est la fonction de répartition (distribution) de la loi normale centrée réduite.
- $\mathbb{P}$  la probabilité historique.
- $\mathbb{Q}$  la probabilité risque-neutre.

---

6. 1 - Nous utilisons les notations de l'article [?].

- $X(\omega_i)$  la  $i^{\text{ème}}$  réalisation de la variable aléatoire  $X$  sous la probabilité historique.
- $X(\omega_i)$  la  $i^{\text{ème}}$  réalisation de la variable aléatoire  $X$  sous la probabilité risque-neutre.
- $\mathbb{Q}^{R_i}$  la probabilité risque neutre associé au monde choqué associé au risque  $R_i$
  
- $N_{MC}$  le nombre de simulations Monte Carlo.
- $N_P$  le nombre de simulations primaires.
- $N_S$  le nombre de simulations primaires.



# Chapitre 1

## Générateur de Scénarios Economiques

DANS CE CHAPITRE, NOUS ALLONS ESSAYER DE RÉPONDRE AUX QUESTIONS

- QU'EST CE QU'UN GÉNÉRATEUR DE SCÉNARIOS ÉCONOMIQUES ?
- POURQUOI A-T-ON BESOIN D'UN GSE DANS L'INDUSTRIE DE L'ASSURANCE ?
- QUELLES SONT LES ÉTAPES DE CONSTRUCTION D'UN GSE ?

## 1.1 Introduction

Commençons par la définition d'un **contrat d'assurance** :

Le contrat d'assurance est un contrat générateur d'engagements réciproques de l'**assuré** et de l'**assureur**. En contrepartie du versement d'une **prime**, aussi appelée cotisation, l'entreprise d'assurance garantit à l'assuré des **prestations** pécuniaires (indemnités ou prestations forfaitaires) ou en nature (assistance juridique, ...) **dans le cas où** le risque prévu dans le contrat **se réaliserait** (vol, accident, ...).

Cette définition rend l'industrie de l'assurance singulière. En effet, dans les industries classiques, le prix d'achat (prix des matières premières) est connu à l'avance avant le prix de vente. Dans l'industrie de l'assurance, il y a **inversion du cycle de production** c'est-à-dire l'entreprise d'assurance doit établir une tarification *judicieuse* d'un produit avant de connaître son coût exact (réalisation de l'aléa). Les primes sont calculées alors que les engagements ne sont pas précisément connus.

Cette vision modifie la lecture du bilan. Dans l'entreprise classique on s'intéresse à la façon dont l'actif (immobilisation, stock) a été financé par le passif (dette). Dans l'entreprise d'assurance, on examine plutôt la façon dont le passif (**provisions techniques** = engagements de l'entreprise d'assurance), a été **couvert** par l'actif (**titres financiers** = investissement des primes d'assurances reçues). On peut considérer que la vente d'un contrat d'assurance s'assimile à la vente d'une option de vente, l'assuré détenant le droit de vendre un actif dont la valeur est garantie.

Ainsi, l'entreprise d'assurance doit pouvoir développer une gestion de risque qui permette de mesurer au mieux

- les risques **assurantiels** (vie et/ou non vie) souscrits auprès des assurés
- les risques **financiers** portés par les actifs du bilan
- leurs **interactions**

afin d'anticiper la rentabilité de son activité et le respect de ses engagements envers ses assurés.

Ces risques techniques ou financiers dépendent de facteurs qui **évoluent** dans le **temps**. Une vision **prospective** devient indispensable, vision qui est consacré par les approches du CFO forum<sup>1</sup> et par la directive européenne **Solvabilité II**.

Dans ce cadre, pour les entreprises d'assurances, afin de

- **déterminer** leur politique **commerciale**<sup>2</sup>
- **maîtriser** leurs **risques** permettant de **pérenniser leur activité**

se doter d'un **outil de simulation** permettant de générer des trajectoires **futures** des différents facteurs de risques pour les intégrer dans les modèles de gestion devient **incontournable**.

---

1. Groupe de discussion de groupes d'assurance européen.

2. Et politique de souscription, ...

Cet outil c'est le **Générateur de Scénarios Economiques** (GSE<sup>3</sup>).

F. Planchet et al. proposent dans leur livre "Scénarios économiques en assurance" [52] la définition globale suivante :

"*Un scénario économique correspond à une projection de grandeurs **économiques** et **financières** sur un horizon d'intérêt*" telles que :

- les rendements d'indices "Action"
- les taux d'intérêt pour les différentes maturités
- les spreads de crédit pour les obligations d'état et privées
- les rendements d'indices "Immobilier"
- le taux d'inflation

Les entreprises d'assurance s'intéressent à l'**évolution** de ces grandeurs financières car elles **conditionnent** le prix des **actifs** financiers de leur bilan et indirectement l'évaluation de leur **passif** via les **interactions** actif-passif. Ainsi, une variation de prix futurs des actifs a un impact sur la richesse de l'entreprise d'assurance.

Un **Générateur de Scénarios Economiques** (GSE) est donc l'outil de modélisation permettant de **générer** de multiples **trajectoires** de ces grandeurs sur un **horizon** donné.

Mais le terme générique GSE peut couvrir en fait **plusieurs concepts** mathématiques plus ou moins complexe :

- Techniques de simulation
- Estimations statistiques
- Intégration d'événement extrêmes
- Structures de dépendance

qui feront l'objet des chapitres suivants.

D'autre part, la construction d'un GSE au-delà de l'aspect mathématique présente plusieurs difficultés. En effet, il doit être **adapté** à l'usage qui en est fait<sup>4</sup> :

- maîtrise des risques, on aura besoin en particulier de simulations sous probabilité historique

---

3. ESG - Economic Scenario Generator.

4. C'est l'objet de la section "Usages du GSE\* en assurance" page 25.

- valorisation, on aura besoin de simulations sous probabilité risque-neutre

Le générateur que **nous** allons développer dans ce mémoire s'inscrit dans le cadre de simulations sous probabilité historique.

Avant de rentrer dans le vif du sujet, revenons sur les notions de de probabilité **historique** et de probabilités **risque-neutre** sous lesquelles les trajectoires seront générées.

### 1.1.1 Simulations sous probabilité historique

La valeur future des grandeurs financières est **aléatoire**, mais lorsque l'on étudie la distribution **historique effective**, on observe pour chaque grandeur des structures statistiques **caractéristiques**. Par exemple, les distributions marginales des grandeurs étudiées peuvent présenter

- une **leptokurticité**<sup>5</sup> : les queues de distributions observées sont plus **épaisses** que celles de la loi normale ou log-normale.
- une **asymétrie**<sup>6</sup> : la probabilité de rendements **négatifs** est supérieure à la probabilité de rendements **positifs**.
- ...

Et les **dynamiques** observées des données historiques peuvent présenter par exemple :

- une tendance
- un retour à la moyenne<sup>7</sup>
- une **dépendance de queues**<sup>8</sup> variable : les dépendances sont **plus faibles** dans les conditions de marché **normal** que dans les conditions de marché **stressé**.
- un phénomène d'**hétéroscédasticité** : les volatilités élevées<sup>9</sup> ont tendance à se regrouper dans le temps.
- ...

Dans le cadre de la théorie du choix dans l'incertain<sup>10</sup>, on suppose que les événements **futurs** possibles sont associés à des probabilités **objectives** basées sur ces caractéristiques observées sur lesquelles les agents économiques s'**accordent**.

En revanche, rien ne garantit que le phénomène **réel**<sup>11</sup> est exactement représenté par ces probabilités objectives.

---

5. Heavy tails - Kurtosis Cf. Coefficient d'aplatissement.

6. Skewness Cf. Coefficient de symétrie.

7. Mean-reversion.

8. Tail dependance.

9. Volatility cluster.

10. Cf. la notion de **loterie** en micro-économie de l'incertain.

11. Ainsi l'utilisation de la terminologie souvent utilisée de **monde-réel** (*real world* en anglais) n'est pas forcément adéquate.

Cette probabilité **objective** est souvent appelé probabilité **historique** que l'on notera  $\mathbb{P}$ .

Notons que simuler sous probabilité historique **n'**implique **pas** de reproduire pas l'historique lui-même. La distribution objective n'est pas la distribution empirique.

D'autre part, on suppose l'**aversion**<sup>12</sup> aux risques des agents économiques. Ainsi, les prix des actifs intègrent une prime de risque pour compenser une perte de rentabilité due aux facteurs de risque sous-jacents.

Par exemple, pour les produits de taux tels que les obligations, les facteurs de risques sont :

- le risque de taux<sup>13</sup>
- le risque de crédit compensé par la prime (*spread*) de crédit
- le risque de liquidité compensé par la prime de **illiquidité**.
- le risque de corrélation compensé par la prime de corrélation pour compenser le risque systémique de défaut sur plusieurs émetteurs.

Pour les actions les facteurs de risques sont

- le risque sur l'évolution du dividende
- le risque sur le ratio dividende/prix<sup>14</sup>

Nous voyons que les facteurs de risque **diffèrent** d'un actif à l'autre. Les actifs n'offrent donc **pas** le **même** rendement **espéré**. Ainsi le rendement d'un actif se décompose comme la somme du rendement sans risque et de la somme des primes de risque associés à chacun des facteurs de risques :

$$\mathbb{E}[r_t^i] = \mathbb{E}[r_t^0] + \sum_{j(i)} \lambda_t^j$$

Par conséquent, toutes les actualisations de flux futurs doivent être faites à un taux intégrant cette prime de risque.

Pour pouvoir **projeter** une situation **économique** déterminée par les différentes variables financières, le **générateur de scénarios économiques** sous probabilité **historique** constituera l'outil de simulation permettant de générer des trajectoires **réalistes**<sup>15</sup> c'est-à-dire permettant de retrouver dans les valeurs futures les propriétés statistiques des marginales observées et les primes de risques associées.

Nous parlerons de **scénarios historiques**, les trajectoires générés par le GSE sous probabilité historique.

---

12. Cf. les travaux de von Neumann et Morgenstern.

13. le produit de taux **sans risque** est supposé soumis à ce seul risque.

14. Earnings per share.

15. Cf. par exemple le paragraphe "Calibration" du document du CEIOPS [70].

En pratique, les primes de risque sont très **difficiles** à déterminer et à modéliser du fait qu'elles soient très différentes d'un actif à l'autre. Une sous-estimation de ces primes de risque revient à sur-estimer la valorisation. Ainsi on introduit la probabilité **risque-neutre**. C'est l'objet du paragraphe suivant.

### 1.1.2 Valorisation sous probabilité risque-neutre

Valoriser sous probabilité risque-neutre consiste à considérer les agents économiques **neutre** à l'égard du risque c'est-à-dire on considère que les primes de risques sont toutes **nulles**<sup>16</sup>. Les actifs ont la **même** espérance de rentabilité (rendement), celui du taux sans risque.

Les prix des produits financiers peuvent être ainsi calculés plus simplement à l'aide de l'espérance des flux futurs (payoffs) actualisés au **taux sans risque** sans avoir à estimer les primes de risque des actifs sous-jacents.

La valorisation des produits financiers revient donc souvent à calculer par **Monte Carlo** une espérance conditionnelle sous la probabilité risque-neutre.

Nous parlerons de **scénarios risque-neutres**, les trajectoires générées par le GSE sous la probabilité **risque-neutre**  $\mathbb{Q}$ . Dans le cadre théorique, les dynamiques des prix des actifs dépendent de paramètres. On **calibre**<sup>17</sup> ces paramètres sur les prix observés.

Dans cette approche, il faut noter que la probabilité risque-neutre **ne** "modifie" **pas** les événements c'est-à-dire que les événements négligeables dans l'un des mondes sont également négligeables dans l'autre monde. On parle alors d'**équivalence** entre les probabilités.

L'**unicité** de la probabilité risque-neutre et l'**équivalence** entre les probabilités historique et risque-neutre suppose que des hypothèses d'**absence d'opportunité d'arbitrage** et de **complétude** du marché soient vérifiées (coût de transactions nuls, liquidité, ...). C'est l'objet des paragraphes suivants.

#### 1.1.2.1 Valorisation en AOA

Nous venons d'introduire les notions de probabilité historique et risque-neutre, introduisons un exemple<sup>18\*</sup> simple permettant d'illustrer leur utilisation.

##### Exemple 1.1.1 (Jeu 1)

*Considérons un jeu de pile ou face équilibré. La mise est de 1 euro et le tirage aléatoire permette d'obtenir un gain d'un euro avec une probabilité de 0,5 ou une perte d'un euro avec une probabilité de 0,5 euros.*

*L'espérance de gain est de*

$$0,5 \times 1 + 0,5 \times -1 = 0$$

*Ici la probabilité historique est de 0,5.*

*Si le joueur est neutre au risque le prix de ce jeu doit être égal à son espérance de gain.*

---

16. Propriété de martingales.

17. Market-consistent calibration.

18. Cf. l'illustration de la présentation [60].

Supposons maintenant que le prix de ce ne soit pas nul, et qu'il existe donc une prime de risque. Le joueur risquophile devra payer par exemple  $P_1 = 10$  centimes.

Le monde risque-neutre consiste à **modifier** les probabilités historiques tel que l'espérance de gain soit égal au prix.

$$0,55 \times 1 + 0,45 \times -1 = 0,1$$

### Exemple 1.1.2 (Jeu 2)

Considérons maintenant un jeu de pile ou face tel que la mise soit également de 1 euro mais tel que le tirage aléatoire permettent soit d'obtenir un gain d'un euro ou soit de récupérer sa mise d'un euro.

L'espérance de gain est de

$$0,5 \times 1 + 0,5 \times 0 = 0,5$$

Ici la probabilité de 0,5 est toujours la probabilité historique.

### Exemple 1.1.3 (Stratégie)

Considérons maintenant la stratégie suivante :

- achat de deux jeux 1
- vente d'un jeu 2

Le gain de cette stratégie n'est plus aléatoire. Dans tous les cas on a :

$$2 \times 1 - 1 \times 1 = 1$$

ou

$$2 \times 0 - 1 \times -1 = 1$$

avec une probabilité de 0,5.

Sachant le prix  $P_1$  du jeu 1, le prix viable<sup>19</sup> du jeu 2 est déterminé par :

$$2 \times P_2 - P_1 = 1 \Leftrightarrow P_2 = \frac{1 + P_1}{2} = 0,55$$

Et si l'on utilise la probabilité risque-neutre on trouve bien

$$0,55 \times 1 + 0,45 \times 0 = 0,55$$

Cela exprime la neutralité au risque.

Cet exemple permet d'illustrer le fait que la probabilité **risque neutre** ne peut pas servir pour calculer l'espérance de gain et **ne permet** pas donc **pas** une approche de **gestion des risques** qui utilise plutôt la probabilité **historique**. La probabilité risque-neutre quant à elle permet plutôt de calculer des **prix** viables.

---

19. i.e. sans possibilité d'arbitrages. On parle de prix en Absence d'Opportunité d'Arbitrage (AOA).

### 1.1.2.2 Marché complet

Dans le cadre d'un marché **complet**, il est possible de couvrir **tout** flux financier, par l'actif **sans risque** et des actifs **risqués**.

En AOA<sup>20</sup> et en marché complet, l'**unique** mesure-martingale est la probabilité **risque-neutre** permettant d'obtenir un prix viable.

En assurance, ce prix en  $t = 0$  s'exprime souvent comme l'espérance sous la mesure produit  $\mathbb{P}^a \otimes \mathbb{Q}^f$  :

$$\boxed{\mathbb{E}^{\mathbb{P}^a \otimes \mathbb{Q}^f} \left[ \sum_{t>0} e^{-\int_0^t r_s ds} F_t \mid \mathcal{G}_0 \right]} \quad (1.1)$$

où

- $\mathbb{P}^a$  est la probabilité décrivant la distribution des aléas **non-financiers** d'assurance.
- $\mathbb{Q}^f$  est la probabilité décrivant la distribution (risque-neutre) des aléas **financiers**.
- $r_t$  le taux d'intérêt instantané **sans risque**.
- $\delta_\bullet = e^{-\int_0^\bullet r_s ds}$  représente le processus d'**actualisation**.
- $F_t$  le flux assurantiel aléatoire **future** à la date  $t$ .
- $\mathcal{G}_t$  représente la tribu **grossière**  $\mathcal{G}_t := \mathcal{F}_t \vee \mathcal{A}_t$  où  $\mathcal{F}_t$  représente l'information financière avant la date  $t$  et  $\mathcal{A}_t$  représente l'information assurantiel (biométrique<sup>21</sup>) avant la date  $t$ .

Dans le cas où le flux assurantiel  $F_t$  **ne** dépend **pas** d'aléas financiers et c'est le cas lorsqu'il **n'y a pas** d'interaction actif-passif<sup>22</sup>, l'équation (1.1) se **simplifie** :

$$\begin{aligned} \mathbb{E}^{\mathbb{P}^a \otimes \mathbb{Q}^f} \left[ \sum_{t>0} e^{-\int_0^t r_s ds} F_t \mid \mathcal{G}_0 \right] &= \sum_{t>0} \mathbb{E}^{\mathbb{Q}^f} [\delta_t \mid \mathcal{F}_0] \times \mathbb{E}^{\mathbb{P}^a} [F_t \mid \mathcal{A}_0] \\ &= \sum_{t>0} P(0, t) \mathbb{E}^{\mathbb{P}^a} [F_t \mid \mathcal{A}_0] \end{aligned} \quad (1.2)$$

On retrouve ainsi la formule standard de la valorisation probabilité sous risque-neutre mais dont le flux est donné par  $\mathbb{E}^{\mathbb{P}^a} [F_t \mid \mathcal{A}_0]$ .

Or souvent le flux  $F_t$  est conséquence d'**interactions** actif-passif, c'est le cas en assurance-vie. Il n'y a donc pas indépendance. On utilise alors souvent la simplification de l'espérance (1.1) en séparant les calculs d'espérance des différentes probabilités :

$$\boxed{\mathbb{E}^{\mathbb{Q}^f} \left[ \mathbb{E}^{\mathbb{P}^a} \left[ \sum_{t>0} e^{-\int_0^t r_s ds} F_t \mid \mathcal{A}_0 \right] \mid \mathcal{F}_0 \right]} \quad (1.3)$$

Le calcul se résume donc à évaluer l'espérance, via la méthode **Monte-Carlo** sous la probabilité **risque-neutre**  $\mathbb{Q}^f$  **sachant** la réalisation des aléas **assurantiels**<sup>23</sup>  $\mathcal{A}_0$ .

20. Cf. les travaux de Harisson, Krepes et Pliska - Pour une vision historique Cf. l'article [11].

21. En assurance-vie.

22. Cf. le paragraphe sur la gestion actif-passif (1.2.3.7).

23. Par exemple en assurance Vie les réalisations sont évaluées via les tables de mortalité.



## 1.2 Usages du GSE en assurance

Le besoin de développer des GSE pour les entreprises d'assurance<sup>24</sup> s'est faite ces dernières années sous l'impulsion d'évolutions réglementaires (Solvabilité II<sup>25</sup>) et comptables (IFRS 4<sup>26</sup>) et de façon parallèle pour des besoins de communication financière (CFO Forum)<sup>27</sup>. Dans ce mémoire, nous aborderons<sup>28</sup> surtout la partie Solvabilité II.

Ainsi dans ce cadre, les entreprises d'assurances ont multiplié des outils de simulation et d'aide à la décision afin de répondre à leurs problématiques de :

- pilotage financier.
- stratégie d'allocation d'actifs.
- gestion actif-passif.
- calcul de la solvabilité globale qui repose sur la notion de provision Best Estimate.

Néanmoins ces outils de simulation doivent être construit de façon *pertinente et judicieuse*. En effet

1. le choix de la probabilité dépend de la nature de calcul de la problématique assurantielle :
  - La probabilité **risque-neutre** est utilisée dès que l'approche financière de valorisation en **valeur de marché**<sup>29</sup> est requise.
  - La probabilité **historique** est utilisée si l'on souhaite obtenir dans une perspective **prospective** des valeurs probables des grandeurs économiques.
2. les simulations sont **chronophage** et le temps de calcul est un élément critique pour les assureurs.

Plus singulièrement, il s'agit d'estimer au mieux les **distributions** futures des variables d'intérêt plutôt que d'obtenir une **prévision** précise.

Dans les paragraphes suivants nous revenons sur le cadre prudentiel de Solvabilité II et les approches d'évaluation des entreprises d'assurance afin d'illustrer l'usage des différentes simulations du GSE. Nous insisterons à chaque fois lorsqu'il s'agira de simulations sous probabilité historique ou de simulations risque-neutre.

Enfin nous présenterons la gestion actif-passif en assurance-vie, du fait des interactions actif-passif rend les calculs actuariels utilisant le GSE plus complexes.

---

24. Ainsi que les entreprises de réassurance et les fonds de retraite.

25. Solvency II Cf. [69].

26. Cf. indiquer une référence.

27. Cf. [73].

28. Néanmoins, on remarque une convergence des différentes notions.

29. Market consistent.

## 1.2.1 Valorisation économique

Sous **Solvabilité I**, les fonds propres des entreprises d'assurance représentent la richesse **accumulée**. Solvabilité I s'inscrit dans une approche **coût historique**.

Afin de valoriser de façon plus *réaliste* les entreprises d'assurance, le CFO Forum et la directive européenne **Solvabilité II**<sup>30\*</sup> ont développé une approche permettant d'incorporer la richesse **future** dans l'évaluation des fonds propres des entreprises d'assurances.

### 1.2.1.1 Bilan économique au sens Solvabilité II

Le considérant 45 de la directive [69] de Solvabilité II considère que "*l'évaluation de la situation financière des entreprises d'assurance et de réassurance devrait reposer sur des principes économiques sains et faire un usage optimal des informations fournies par les marchés financiers ainsi que des données généralement disponibles sur les risques techniques en matière d'assurance. En particulier, les exigences de solvabilité devraient reposer sur une évaluation économique fondée sur le total du bilan.*".

Ce principe d'évaluation économique des bilans des entreprises d'assurance a pour but de refléter au *mieux* leur position financière :

- L'actif sera valorisé en **valeur de marché**<sup>31</sup>.
- Le passif sera évalué de **façon cohérente avec le marché**<sup>32</sup>.

La valorisation économique du passif correspond au montant d'actifs qu'un tiers exigerait pour reprendre les engagements de l'entreprise d'assurance.

Nous n'utilisons pas le vocable valeur de marché pour le passif car il n'existe pas vraiment de marché **co-tant** les passifs. Si le passif est **réplicable** c'est-à-dire que les flux des passifs peuvent être répliqués pas des produits financiers alors la valeur économique du passif sera la valeur de marché de ces produits financiers. Dans la majorité des cas, les passifs ne sont pas répliquables, on introduit alors la notion de **Best Estimate** et de **marge de risque**.

L'article 77 de la directive européenne [69] définit les **provisions techniques** comme la "somme de la meilleure estimation, **Best Estimate**, et de la **Marge de Risque**" :

La **meilleure estimation** correspond à "la **moyenne** pondérée par leur **probabilité** des flux de trésorerie **futurs**, compte tenu de la valeur temporelle de l'argent (valeur **actuelle** attendue des flux de trésorerie futurs), estimée sur la base de la courbe des taux sans risque pertinents" :

$$L_t := \mathbb{E}^{\mathbb{Q}} \left[ \sum_{u>t} \frac{\delta_u}{\delta_t} F_u \mid \mathcal{F}_t^{RW} \right] \quad (1.4)$$

Ici on considère que les flux aléatoires  $F_t$  ont déjà été *probabilisés* par le risque assurantiel. Cf. (1.3) page 24.

---

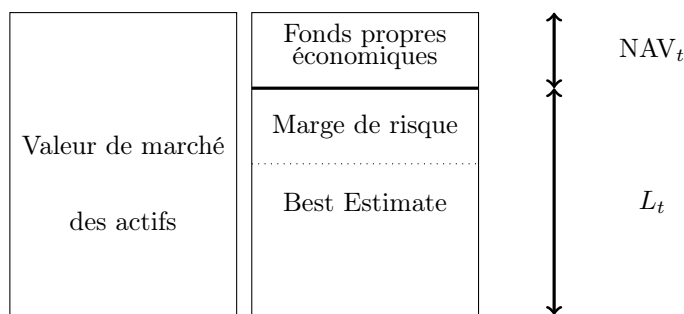
30. Mise en application au 1er janvier 2016.

31. Marked-to-Market.

32. Fair value.

La **marge de risque** est "calculée de manière à garantir que la valeur des **provisions techniques** est équivalente au **montant** que les entreprises d'assurance et de réassurance demanderaient pour **reprendre** et honorer les engagements d'assurance et de réassurance".

Le **bilan prudentiel** se représente donc de la façon suivante :



On appellera **actif net**<sup>33</sup> ou fonds propres **économiques**, la différence entre l'actif économique et les provisions techniques économiques<sup>34</sup> :

$$\text{NAV}_t = A_t - L_t \quad (1.5)$$

### 1.2.1.2 Valeur intrinsèque MCEV

Les entreprises d'assurances publient dans le cadre de leur communication financière la valeur de leur entreprise. Cette valeur peut être utilisée dans le cadre d'un projet de fusion-acquisition par exemple.

La MCEV<sup>35</sup> mesure la valeur de l'entreprise d'assurance du point de vue de l'**actionnaire**. Cette valeur **intrinsèque**<sup>36</sup> **économique** est définie des éléments suivants :

- l'**Actif Net Réévalué** (ANR) représente les richesses **accumulées** par la compagnie d'assurance. Il s'agit de la valeur de marché des actifs adossés aux fonds propres i.e. qui ne couvrent pas les engagements. Il y a prise en compte des plus ou moins values latentes.
- la **Value-In-Force** (VIF) correspond à la valeur **actuelle** des profits **futurs** générés par le portefeuille de contrats en tenant compte des **options** et garanties financières **implicites** :

$$\text{VIF}_t := \mathbb{E}^{\mathbb{Q}} \left[ \sum_{u>t} \frac{\delta_u}{\delta_t} R_u \mid \mathcal{F}_t^{RW} \right] \quad (1.6)$$

Ainsi l'**actif net** peut être défini également comme la somme de la richesse **actuelle** et de la richesse **future** actualisée :

$$\text{NAV}_t = \text{ANR}_t + \text{VIF}_t \quad (1.7)$$

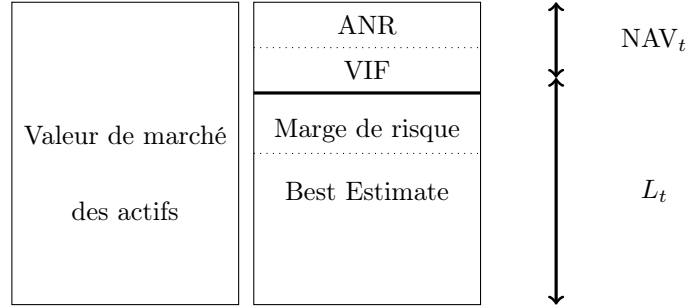
33. Net Asset Value.(NAV

34. Best Estimate of Liabilities - BEL ou BE.

35. Market Consistent Embedded Value. Cf. [73].

36. Embedded value.

Ainsi dans l'approche MCEV, le bilan prudentiel se décompose :



Les formulations (1.4) page 26 et (1.6) page 27 des quantités  $NAV_t$ ,  $VIF_t$  sont approximées par la méthode de Monte Carlo :

$$L_t \simeq \frac{1}{N_{MC}} \sum_{S=1}^{N_{MC}} \sum_{u>t} \frac{\delta_u^S}{\delta_t} F_t^S$$

$$VIF_t \simeq \frac{1}{N_{MC}} \sum_{S=1}^{N_{MC}} \sum_{u>t} \frac{\delta_u^S}{\delta_t} R_t^S$$

où  $R_t^S$  et  $F_t^S$  sont des **réalisations** de trajectoires risque-neutres.

Ainsi le calcul de la VIF et de la valeur économique du passif à la date initiale ( $t=0$ ) nécessitent **seulement** des simulations **risque-neutre** pour déterminer les espérances par Monte Carlo. Dans ce cas les simulations sous probabilité historique ne sont pas nécessaires.

## 1.2.2 Calcul du SCR

La directive Solvabilité II<sup>37</sup> exige que "*les entreprises d'assurance et de réassurance détiennent des fonds propres éligibles couvrant le capital de solvabilité requis*".

L'objet de ce paragraphe est de présenter le calcul du **capital solvabilité requis**.

### 1.2.2.1 Capital économique

Afin d'assurer un niveau de faillite faible, la directive Solvabilité II **exige** un niveau de fonds propre économique issu du bilan économique permettant de couvrir les risques financiers et assurantiels à **un** an avec un niveau de confiance de **99,5%** :

$$\operatorname{argmin}_{x \in \mathbb{R}^+} \mathbb{P}\{\text{NAV}_1 < 0 \mid \text{NAV}_0 = x\} \leq 0.5\% \quad (1.8)$$

Il s'agit donc de calculer<sup>38</sup> pour les entreprises d'assurance un **quantile** avec un niveau de confiance à 99,5% de la **distribution** des fonds propres économiques à **un** an.

37. Cf. Section 4 de la directive européenne [69].

38. Article 101-3. de [69] "*Le capital de solvabilité requis correspond à la valeur en risque (Value-at-Risk) des fonds propres de base de l'entreprise d'assurance ou de réassurance, avec un niveau de confiance de 99,5% à l'horizon d'un an.*"

Néanmoins, il existe plusieurs interprétations<sup>39</sup> de ce calcul mais nous retenons la définition suivante :

On appelle **Capital Economique théorique**, qu'on notera SCR, la solution **approchée** de l'équation (1.8) :

$$\boxed{\text{SCR}_0 = \text{NAV}_0 - P(0, 1) q_{0.5\%}(\text{NAV}_1)} \quad (1.9)$$

avec  $P(0,1)$  le prix d'un zéro-coupon de maturité un an.

La quantité  $-P(0, 1) q_{0.5\%}(\text{NAV}_1)$  représente un montant de fonds propres qu'il est nécessaire d'**ajouter** aux fonds propres **initiaux**  $\text{NAV}_0$  pour **garantir** la condition de non-faillite de l'entreprise d'assurance avec un niveau de confiance de 99.5%.

$q_{0.5\%}(\text{NAV}_1)$  peut être négatif par exemple dans un scénario de baisse importante de la valeur de marché des actifs ( $\text{NAV}_t = A_t - L_t$ ).

Posons  $K := P(0, 1) q_{0.5\%}(\text{NAV}_1)$ . Notons  $\text{NAV}_1^K$ ,  $A_1^K$ ,  $L_1^K$ , l'actif net, la valeur de marché de l'actif et  $L_1$  la provision technique économique à un an **après** ajout du capital  $K$ .

On a

$$\mathbb{P}\{\text{NAV}_1^K < 0 \mid \text{NAV}_0 + K\} = \mathbb{P}\{A_1^K - L_1^K < 0 \mid \text{NAV}_0 + K\}$$

Si les fonds propres sont **cantonés** i.e. il n'y pas d'interaction actif-passif (participation aux bénéfices des assurés en assurance Vie) on a alors  $L_1^K = L_1$ . Ainsi en capitalisant la somme  $K$  sur un an, on obtient :

$$\begin{aligned} \mathbb{P}\{\text{NAV}_1^K < 0 \mid \text{NAV}_0 + K\} &= \mathbb{P}\{A_1^K - L_1 < 0 \mid \text{NAV}_0 + K\} \\ &= \mathbb{P}\{A_1 + \frac{K}{P(0, 1)} - L_1 < 0 \mid \text{NAV}_0 + K\} \\ &= \mathbb{P}\{\text{NAV}_1 - q_{0.5\%}(\text{NAV}_1) < 0 \mid \text{NAV}_0 + K\} \\ &= \mathbb{P}\{\text{NAV}_1 < q_{0.5\%}(\text{NAV}_1) \mid \text{NAV}_0 + K\} \\ &= 0.5\% \end{aligned}$$

Les quantités  $\text{NAV}_1$  et  $L_1$  sont **aléatoires**.

### 1.2.2.2 Approche standard

Dans le cadre du pilier I de Solvabilité II, une formule **standard** (*simple*) pour le calcul du capital de solvabilité requis de base a été proposée<sup>40</sup>. Elle consiste à évaluer un SCR global par **agrégation** de SCR **élémentaires** relatifs à chaque type de **risque**  $R$ , c'est l'approche **modulaire** :

- Risque de marché
- Risque de défaut<sup>41</sup>

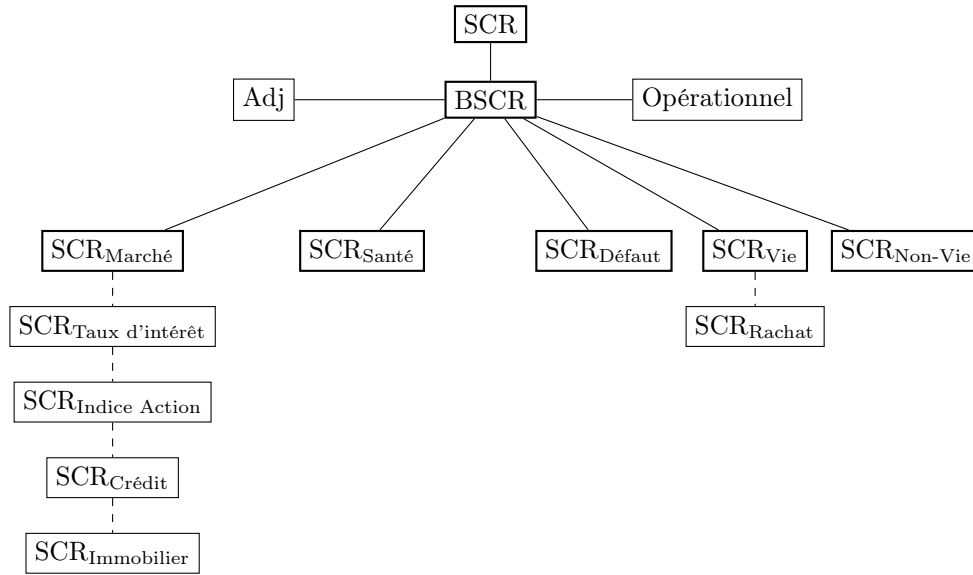
39. Cf. la discussion au paragraphe 2.2 page 37 du livre de Guibert et al. [54].

40. Cf. Article 104 de la directive européenne [69].

41. Risque de contrepartie.

- Risque de souscription Vie
- Risque de souscription en Santé
- Risque de souscription Non Vie

Ci-dessous le schéma succinct<sup>42</sup> de la **pieuvre**, qui définit l'approche standard dit **modulaire**<sup>43</sup>.



Les  $SCR_i$  **élémentaires** sont calculés à partir de chocs **instantanés** des facteurs de risque  $R_i$  : le capital économique associé au sous-risque  $R_i$  correspond à la **variation parler de l'absence de la marge pour risque dans la formule standard** des fonds propres issus du scénario **central** et des fonds propres **après** réalisation d'un choc instantané du facteur de risque  $R_i$ .

Les fonds propres sur le bilan initial (i.e.  $t=0$ ) après choc instantané du facteur de risque  $R_i$  :

$$NAV_0^{R_i} = ANR_0^{R_i} + \mathbb{E}^{\mathbb{Q}-R_i} \left[ \sum_{t \geq 0} \frac{\delta_t}{\delta_0} R_t^{R_i} \mid \mathcal{F}_0^{R_i} \right]$$

On **approxime** l'espérance conditionnelle par la méthode **Monte Carlo**

$$NAV_0^{R_i} \simeq ANR_0^{R_i} + \frac{1}{N_{MC}} \sum_{S=1}^{N_{MC}} \sum_{t \geq 0} \frac{\delta_t^S}{\delta_0} R_t^{S, R_i}$$

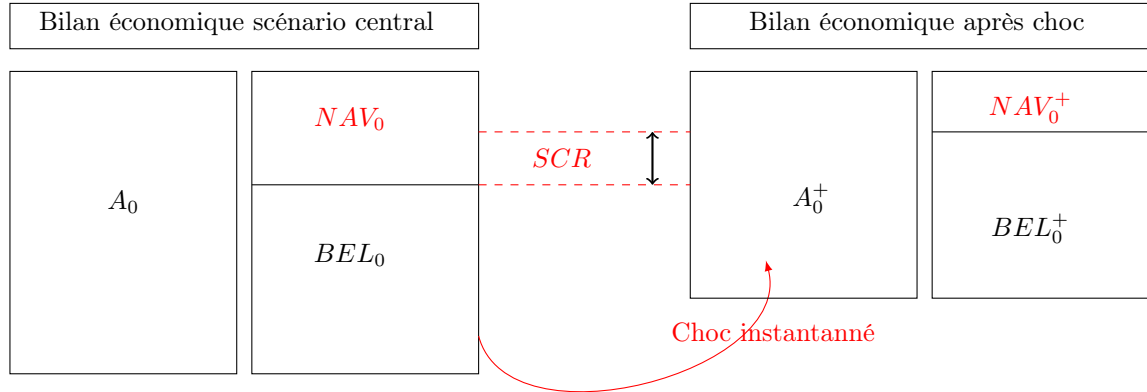
42. Nous nous intéressons dans ce mémoire principalement au risque de marché. Nous faisons apparaître le risque de rachat du fait de l'interaction actif-passif.

43. Bottom-up Approach.

Dans le cadre de la formule standard, les calculs **ne** nécessitent donc **pas** de simulations sous probabilité historique. Seules les simulations **risque-neutres** issues de  $t = 0$  pour évaluer les espérances conditionnelles de chaque sous-module sont nécessaires.

Le **capital économique** associé au risque  $R_i$ , appelé SCR élémentaire, correspond à la **variation** de fonds propres entre le scénario **central** et le scénario **choqué** :

$$\text{SCR}_0^{R_i} := \max\{\text{NAV}_0 - \text{NAV}_0^{R_i}, 0\}$$



Le SCR est obtenu en agrégeant les  $\text{SCR}_0^{R_i}$  élémentaire via des **matrice de corrélation pré-définies**<sup>44</sup>.

Le calcul des SCR **élémentaires ne tient pas compte** des spécificités de chaque entreprise d'assurance. Il existe une variante de la formule standard "Standard Entity Specific" permettant une calibration des chocs plus en adéquation avec l'entreprise d'assurance.

L'introduction des matrices de corrélation pour agréger les SCR de chaque sous-module pour obtenir le SCR total permet de prendre en compte un effet de diversification. La diversification réduit le capital réglementaire.

### 1.2.2.3 Stochastique dans le Stochastique - SdS

Nous venons de présenter la méthode standard qui **ne** correspond **pas** exactement à la définition du capital économique :

$$\text{SCR}_0 = \text{NAV}_0 - P(0, 1) q_{0.05\%}(\text{NAV}_1)$$

Le calcul exact est appelé Stochastique dans le Stochastique<sup>45</sup>. Nous allons expliquer dans ce paragraphe le

44. Les matrices de corrélation inter-risque sont définies dans l'annexe 4 de [69]. Les chocs de chaque sous-risques sont calibrés sur la base d'une mesure de la VaR à 99,5%. Cf. Article 104.4. de [69].

45. On parle également de **Simulations dans les Simulations**. En anglais on parle de **Nested simulations**.

sens de cette expression.

La quantité aléatoire  $NAV_1$  peut être exprimée comme une espérance **conditionnelle** sous la probabilité risque-neutre :

En date  $t=1$ , la valeur de l'actif net en 1 est égale à l'espérance sous probabilité **risque-neutre** des profits futurs ( $t \geq 2$ ) actualisés au taux d'intérêt continu sans risque :

$$NAV_1 = ANR_1 + \mathbb{E}^{\mathbb{Q}}\left[\sum_{t \geq 2} \frac{\delta_t}{\delta_1} R_t \mid \mathcal{F}_1\right] \quad (1.10)$$

En date  $t = 1$ , la valeur du passif est égale à l'espérance sous probabilité **risque-neutre** des flux futurs ( $t \geq 2$ ) actualisés au taux d'intérêt continu sans risque :

$$L_1 = \mathbb{E}^{\mathbb{Q}}\left[\sum_{t \geq 2} \frac{\delta_t}{\delta_1} F_t \mid \mathcal{F}_1\right] \quad (1.11)$$

Nous remarquons que ces deux évaluations risque-neutre dépendent de l'information en date  $t=1$ . Les formulations (1.10) et (1.11) nécessitent <sup>46</sup> donc

- des simulations sous probabilité **historique** pour déterminer la filtration  $\mathcal{F}_1$ .
- des simulations **risque-neutre** pour déterminer par Monte Carlo l'espérance conditionnelle  $\mathbb{E}^{\mathbb{Q}}[\cdot \mid \mathcal{F}_1]$

Il s'agit de simulation **imbriquées**, c'est le **stochastique dans le stochastique**.

Le calcul du SCR s'effectue en deux étapes :

1. Pour un calcul prospectif à  $x$  années on simule des trajectoires en **monde réel** sur  $x$  années. On appellera ces simulations **trajectoires primaires**.
2. Dans un deuxième temps, on simulera des trajectoires dites **secondaires** sous la probabilité *risque-neutre conditionnellement* à chaque trajectoire *primaire* générée sous probabilité historique.

Nous obtenons ainsi une distribution de la variable aléatoire  $NAV_1$  dont la  $i^{\text{ème}}$  réalisation est

$$NAV_1(\omega_i) = ANR_1(\omega_i) + \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} \sum_{t \geq 2} \frac{\delta_t(\omega_i, \tilde{\omega}_j)}{\delta_1(\omega_i)} R_t(\omega_i, \tilde{\omega}_j)$$

On obtient le **capital économique** en prenant le quantile **empirique** d'ordre 0.5% de l'échantillon des  $NAV_1(\omega_i)$  :

$$\boxed{SCR_0 = NAV_0 - P(0, 1) NAV_{1, (\lfloor N_p \times 0.5\% \rfloor + 1)}} \quad (1.12)$$

46. **Quelques fois on n'a pas besoin de simulations à préciser. Redemander à FP**

47. Graphe de l'article [15] - avec les notations  $FP_1 := NAV_1$  et  $VEP_1 := L_1$



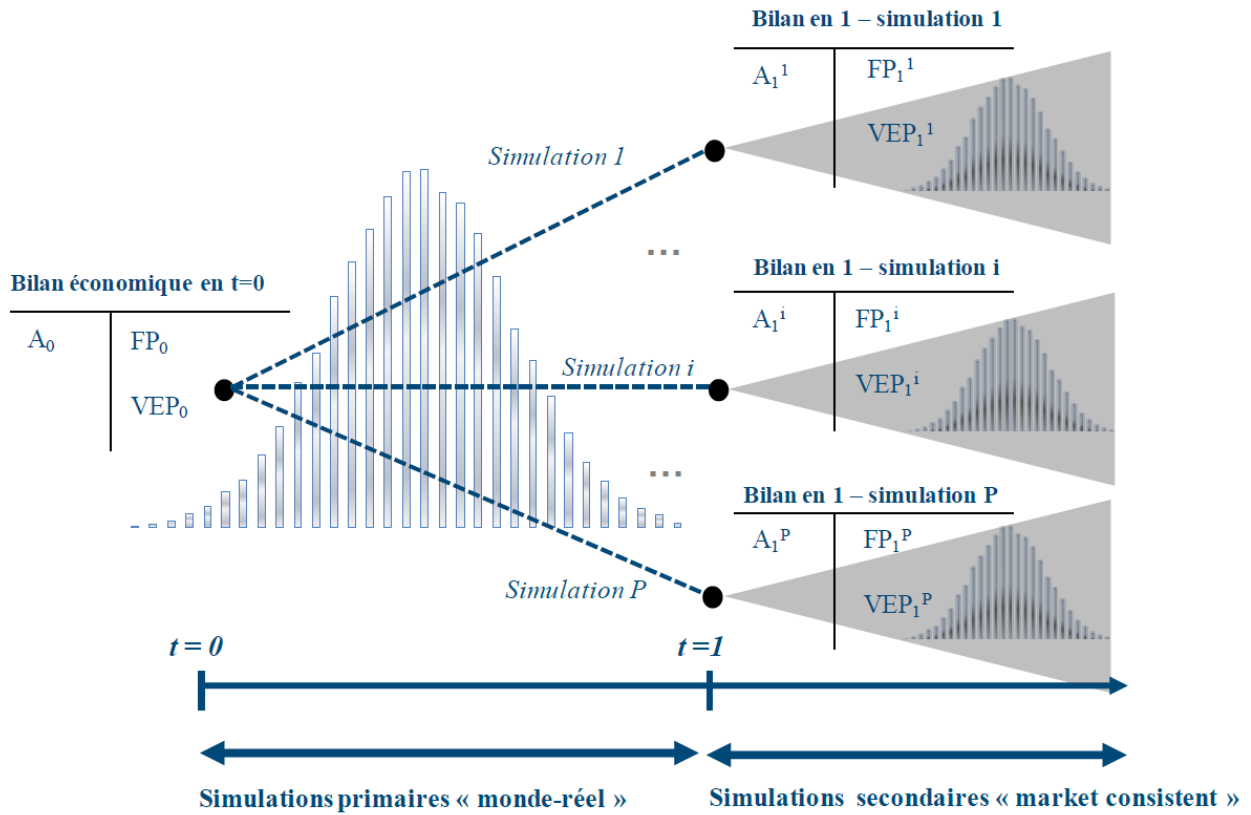


FIGURE 1.1 – Distribution de fonds propres économiques par la méthode SdS<sup>47</sup>

$N_P$  est le nombre de simulations primaires.

Pour  $P$  simulations primaires et  $S$  simulations secondaires le nombre de simulations total est de  $S \times P$ . Celui-ci doit être grand pour pouvoir assurer la convergence des Monte Carlo sous-jacents pour le calcul final du quantile d'ordre 0,5%. Cette approche SdS est **chronophage**, elle induit un temps calcul très long du fait des simulations imbriquées et des interactions actifs-passifs. Il est nécessaire du point de vue **opérationnel** de déterminer des méthodes d'approximation permettant de réduire le nombre de simulations.

Il est possible de diminuer le temps de calcul en utilisant la méthode SdS **seulement** sur certains sous-modules et d'utiliser la méthode **standard** sur les autres sous-modules. Il s'agit du modèle **interne partiel**.

Des méthodes **alternatives** au calcul SdS ont été proposées afin de **réduire** le temps de calcul tout en permettant d'obtenir une estimation robuste du Capital Economique. Elles proposent de contourner le problème de temps de calcul :

- en **diminuant** directement le nombre de simulations primaires et/ou secondaires. Par exemple, les méthodes suivantes :
  - Portefeuille répliquant
  - Curve fitting
  - Least Square Monte Carlo
- en **sélectionnant** certaines trajectoires, celles qui contribuent au calcul du SCR. Par exemple :
  - scénarios de stress test à dire d'expert souvent utilisés en tant que modèle-jouet <sup>48</sup>; par exemple pour évaluer des stratégies de règles de gestion <sup>49\*\*</sup> (ventes d'actifs, ...), des stratégies des couvertures (caps, ...), de la justification de la calibration des chocs des facteurs de risques, ...
  - accélération du SdS

Ces méthodes alternatives feront l'objet des paragraphes suivants.

#### 1.2.2.4 Portefeuille répliquant

Lorsqu'on souhaite valoriser le passif **économiquement** c'est-à-dire de façon cohérente avec le marché, il devient naturel de vouloir utiliser les outils de **finance mathématique**.

De plus, en particulier en assurance Vie, les passifs comportent des **garanties** qui peuvent être assimilées à des **options** financières (*caps, floors, swaptions, ...*) :

- Taux à Minimum Garanti
- Participation aux bénéfices
- Rachat conjoncturel
- Arbitrage entre fonds UC et fonds euros
- Option de conversion de rente
- Garantie plancher pour les UC en cas de décès

L'objet des méthodes de portefeuilles répliquants est donc de **répliquer le passif** par des **instruments financiers** et de valoriser en valeur de marché ces produits pour connaître la valeur économique du passif.

L'hypothèse **implicite** de la méthodologie des portefeuilles répliquants est donc l'**adéquation** des flux de passifs à ceux du portefeuille répliquant : i.e. le *pay-off*  $F_t$  en  $t$  du passif se décompose en  $k$  *pay-offs* en  $t$

---

<sup>48.</sup> toy-model

<sup>49.</sup> Management rules

d'instruments financiers  $h_t^j$  élémentaires :

$$F_t := \sum_k \alpha_k h_t^k \quad , \hat{\alpha}_j \in [0, 1] \quad (1.13)$$

On détermine les poids  $\alpha_k$  selon un critère de **moindres carrés**. Deux principaux critères sont utilisés :

— Le premier critère : adéquation des flux<sup>1</sup> à toute date  $t$  :

$$\hat{\alpha}_j := \operatorname{argmin}_{\alpha_k} \sum_{i=1}^S \sum_{t>0} \left( F_t(\tilde{\omega}_i) - \sum_k \alpha_k h_t^k(\tilde{\omega}_i) \right)^2$$

— Le deuxième critère : adéquation des valeurs actuelles nettes<sup>2</sup>

$$\begin{aligned} \hat{\alpha}_j &:= \operatorname{argmin}_{\alpha_k} \sum_{i=1}^S \left( \sum_{t>0} \delta_t(\tilde{\omega}) F_t(\tilde{\omega}_i) - \sum_k \alpha_k \delta_t(\tilde{\omega}) h_t^k(\tilde{\omega}) \right)^2 \\ &= \operatorname{argmin}_{\alpha_k} \sum_{i=1}^S \left( \text{VAN}_0(\tilde{\omega}) - \text{RP}_0(\tilde{\omega}) \right)^2 \end{aligned}$$

où  $S$  est le nombre de simulations risque-neutre<sup>50</sup>.

Dans un modèle de régression linéaire multiple on a

$$\begin{aligned} F_t &\simeq \sum_k \alpha_k h_t^k \\ &= \sum_k \alpha_k h_t^k + \epsilon_t \end{aligned}$$

avec

$$\mathbb{E}[\epsilon_t] = 0$$

On a donc

$$\begin{aligned} L_0 &= \mathbb{E}^{\mathbb{Q}} \left[ \sum_{t>0} \delta_t F_t \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[ \sum_{t>0} \delta_t \left( \sum_k \alpha_k h_t^k + \epsilon_t \right) \right] \\ &= \sum_k \alpha_k \mathbb{E}^{\mathbb{Q}} \left[ \sum_{t>0} \delta_t h_t^k \right] + \sum_{t>0} \mathbb{E}^{\mathbb{Q}} [\delta_t \epsilon_t] \\ &= \text{RP}_0 \end{aligned}$$

---

50. Beaucoup d'auteurs ne précisent pas la probabilité de simulation.

La dernière égalité est vérifiée<sup>51</sup> si les variables  $\delta_t$  et  $\epsilon_t$  sont **décorrélées**.

Une fois les poids déterminés, on peut estimer le Capital Économique en prenant le quantile empirique de la distribution simulée des  $RP_1$  :

$$\widehat{SCR}_0 = NAV_0 - P(0, 1) \widehat{RP}_{1, (\lfloor N_p \times 0.5\% \rfloor + 1)}$$

où  $N_p$  est le nombre de simulations primaires.

Pour certains instruments financiers, les prix de marché peuvent être déterminés par des formules **fermées** sans utiliser des simulations Monte Carlo. Ces instruments  $h^k$  sont à privilégier dans la constitution du portefeuille répliquant afin de réduire le temps de calcul pour déterminer  $RP_1$ .

La valeur de marché du portefeuille de réplcation dans le scénario  $i$  à la date  $t = 1$  est donnée par

$$\widehat{RP}_1(\omega_i) = \sum_k \hat{\alpha}_k h_1^k(\omega_i)$$

Les  $h_1^k(\omega_i)$  sont déterminées par des formules fermées.

L'algorithme de l'estimation du  $SCR_0$  s'effectue donc en trois étapes :

1. Effectuer  $P$  simulations primaires en monde réel.
2. Valoriser par formule fermée les instruments de marché conditionnellement aux  $P$  scénarios primaires.
3. Calculer le quantile empirique d'ordre 0,05% de la distribution simulée  $RP_1$  pour obtenir le Capital Economique.

En conclusion :

La méthodologie de portefeuille **répliquant** permet de réduire le temps calcul. En effet l'utilisation de formules **fermées** dispense de générer des scénarios secondaires, seuls les scénarios primaires sous probabilité **historique** sont **générées**.

### 1.2.2.5 Curve Fitting

Dans le cadre du calcul du Capital Economique (1.12) page 32, la méthodologie du **Curve Fitting** permet de **diminuer** le nombre de simulations.

C'est une méthodologie qui consiste à **identifier** les **facteurs de risques** ( $R_i$ ) sous-jacents au calcul du Capital Economique SCR et d'exprimer la  $NAV_1$  à un an comme un **polynôme** en ses facteurs :

$$NAV_1 = \sum_{\alpha} a_{\alpha} R^{\alpha}$$

---

51. Cf. les remarques sur la garantie de cette hypothèse dans l'article [16].

où  $R^\alpha := R_1^{\alpha_1} \times \dots \times R_n^{\alpha_n}$ .

Une fois les coefficients du polynôme estimés, il est possible de calculer la valeur de la NAV<sub>1</sub> pour **toute** valeur de simulation **primaire** sans devoir effectuer des simulations secondaires. On obtient enfin une distribution empirique de la NAV<sub>1</sub>. On peut donc calculer le quantile empirique d'ordre 0,05%.

L'algorithme s'effectue donc en quatre étapes :

1. Effectuer un SdS pour un nombre de points  $P' < P$ .
2. Estimation du polynôme sur  $P'$  points.
3. Calculer la NAV<sub>1</sub> sur les  $P - P'$  points restants
4. On obtient ainsi une distribution empirique de  $P$  NAV<sub>1</sub>. On calcule enfin le quantile empirique d'ordre 0,05% pour obtenir le Capital Economique.

En conclusion :

Cet algorithme permet donc de **diminuer** le nombre de SdS en effectuant des simulations secondaires sur moins de points de simulations primaires.

### 1.2.2.6 Accélération du SdS

Dans le cadre du Stochastique dans le Stochastique, le calcul du Capital Economique s'intéresse particulièrement à la queue de distribution puisque il s'agit de déterminer un quantile d'ordre 0,5% :

$$SCR_0 = NAV_0 - P(0, 1) NAV_{1, (\lfloor N_P \times 0.5\% \rfloor + 1)}$$

La connaissance de la distribution **complète** à un an n'est donc pas nécessaire.

L'idée de l'algorithme d'accélération du SdS, initié par Devineau et Loisel [15], consiste à déterminer les scénarios **adverse** pour la valeur de la NAV et d'effectuer des simulations **secondaires** pour ces **seuls** scénarios.

L'algorithme d'accélération s'effectue selon les étapes principales suivantes :

1. Déterminer les  $k$  facteurs de risque de l'entreprise d'assurance.
2. Effectuer  $P$  simulations primaires.
3. Calculer la réalisation du  $k$ -uplet de facteurs de risque pour chacune des  $P$  réalisations primaires via le GSE monde réel.
4. Construire une région d'exécution du SdS en associant chaque réalisation du  $k$ -uplet une norme permettant de classer les scénarios primaires et de définir une région d'exécution

en fonction d'un seuil pré-défini. Cette région d'exécution permet de définir les scénarios primaires a priori défavorables pour l'entreprise d'assurance. A ce stade, P' scénarios primaires appartiennent à la région d'exécution.

5. Effectuer pour chacune de P' simulations primaires, S simulations secondaires via le GSE risque-neutre permettant d'évaluer, dans le modèle ALM, P' NAV.
6. Classer par ordre croissant les P' NAV calculées pour les scénarios adverses c'est-à-dire entraînant une NAV faible (on s'intéresse seulement à la queue de distribution pour déterminer le quantile d'ordre 0,5%).
7. Le capital économique correspond au  $(\lfloor N_P \times 0.5\% \rfloor + 1)^{\text{ème}}$  NAV de ce classement.

En conclusion <sup>52</sup> :

Cet algorithme permet donc de **diminuer** le temps de calcul en réduisant le nombre de SdS en effectuant des simulations secondaires risque-neutre **sur moins de points** de simulations primaires en monde réel.

### 1.2.2.7 Least Square Monte Carlo - LSMC

La méthodologie du LSMC <sup>53</sup> a été introduite par Longstaff Schwartz [25] en Finance mathématique pour évaluer les options américaines et puis en assurance <sup>54</sup> pour approximer l'espérance conditionnelle de l'équation de la NAV<sub>t</sub> :

$$\text{NAV}_t = \text{ANR}_t + \mathbb{E}^{\mathbb{Q}} \left[ \sum_{s>t} \frac{\delta_s}{\delta_t} F_s \mid \mathcal{F}_t \right]$$

En effet, si on suppose qu'on peut exprimer la filtration  $\mathcal{F}_t^{RW}$  comme une  $\sigma$ -algèbre générée par des trajectoires **markoviennes** de  $d$  facteurs de risques <sup>55</sup>  $(Y_t)_{t \geq 0} \in \mathbb{R}^d$  alors

$$L_t = \mathbb{E}^{\mathbb{Q}} \left[ \sum_{s>t} \frac{\delta_s}{\delta_t} F_s \mid Y_t, t \in [0, t] \right]$$

Opérationnellement <sup>56</sup>,  $L_t$  ne dépend pas de la trajectoire **complète** des facteurs de risques  $(Y_t, t \in [0, t])$  mais l'information peut être représentée par un nombre **fini** de variable  $D_t := (D_t^1, \dots, D_t^m)$ .

On peut donc ré-écrire l'espérance conditionnelle

$$L_t = \mathbb{E}^{\mathbb{Q}} \left[ \sum_{s>t} \frac{\delta_s}{\delta_t} F_s \mid (Y_t, D_t) \right]$$

52. Préciser la difficulté de construire la région d'exécution pour un nombre de facteurs de risques  $> 2$

53. Certains auteurs utilisent l'acronyme LSM.

54. Cf. l'article de Bauer [7].

55. State Markov process

56. Cf. l'article de Bauer [6].

$L_t$  appartient à l'espace  $L^2(\Omega, \sigma(Y, D), \mathbb{P})$  et on peut donc **décomposer** l'espérance conditionnelle dans une **base** hilbertienne :

$$\begin{aligned} L_t &= \sum_{k=1}^{+\infty} b_k e_k(Y_t, D_t) \\ &\simeq \sum_{k=1}^M b_k e_k(Y_t, D_t) \end{aligned}$$

A priori pour calculer A partir des  $P$  simulations **monde réel**  $Y_1(\omega)$  on obtient une distribution empirique de  $(Y_1^1, D_1), \dots, (Y_1^i, D_1^i), \dots, (Y_1^P, D_1^P)$  et de  $P$  simulations risques neutres

On détermine les coefficients  $b^k$  par un critère de moindres carrés à partir des  $P$  simulations monde réel :

$$\hat{b}_k := \operatorname{argmin}_{b_k} \sum_{i=1}^P \left( \sum_{s>t} \frac{\delta_s(\omega_i, \tilde{\omega}_i)}{\delta_t} F_s(\omega_i, \tilde{\omega}_i) - \sum_{k=1}^M b_k \tilde{e}_k(Y_1(\omega_i, \tilde{\omega}_i), D_1(\omega_i, \tilde{\omega}_i)) \right)^2$$

où  $(\omega_i, \tilde{\omega}_i)$  est **une** trajectoire risque-neutre sur  $\{s > t\}$  conditionnellement à la trajectoire monde réel  $(\omega_i)$ .

Une fois les coefficients déterminés, on peut calculer le Capital Économique :

$$\widehat{\text{SCR}}_0 = \text{NAV}_0 - P(0, 1) \left( \sum_{k=1}^M \hat{b}_k e(Y, D) \right) (\lfloor N_p \times 0.5\% \rfloor + 1)$$

En conclusion :

Cet algorithme permet donc de **diminuer** le nombre de SdS en effectuant une seule des simulation secondaire par simulation primaire.

### 1.2.3 Gestion actif-passif

La gestion actif-passif<sup>57</sup> consiste, pour l'entreprise d'assurance, à assurer une visibilité suffisante de ses résultats futurs en optimisant les emplois et les ressources de son bilan. Il doit donc pouvoir anticiper les impacts sur son bilan.

Par exemple, en assurance Vie, le pilotage doit prendre donc en compte les variations des facteurs de risques (aléas) qui affectent son bilan :

- comportements des assurés rachats (totaux/partiels, structurels/partiels, résiliations, décès, ...)
- participation aux bénéfices
- évolution des actifs (rentabilité, risque marché, risque de crédit, risque liquidité, ...)

afin que toute variation de la valeur de l'actif correspondent à une variation équivalente du passif (ou inversement).

Développer un GSE permet de modéliser l'ensemble des **risques** "marché" susceptibles d'impacter l'actif de l'assureur, et **indirectement** le passif via ce que l'on nommera les **interactions actifs-passifs**.

#### 1.2.3.1 Risque de taux d'intérêt

Les entreprises d'assurance doivent pouvoir quantifier l'impact de variations des taux d'intérêt. En effet, du fait de la structure de leur actif du bilan, ils sont **exposés** au risque de taux d'intérêt, à la **baisse** comme à la **hausse** :

- En cas de **baisse** des taux d'intérêts, l'entreprise d'assurance pourra réaliser des plus-values latentes pour garder une allocation constante du fait de l'appréciation des obligations, ce qui est positif à court terme. Mais à moyen et long terme, si la baisse des taux d'intérêts est trop importante, le **rendement** des **nouveaux** actifs obligataires pourrait être **insuffisant**<sup>58</sup> par rapport aux **engagement**<sup>59</sup> de taux d'intérêt (TMG<sup>60</sup>, effet cliquet) pris envers les assurés.
- En cas de **hausse** des taux d'intérêts, l'entreprise d'assurance subit une **dépréciation** de ses obligations, des moins-values latentes sont alors enregistrées. La concurrence des produits bancaires (Livret A, nouveaux contrats, ...) de taux qui pourrait engendrer des **rachats** et ainsi concrétiser les moins-values latentes (ventes des obligations).

D'autre part, le **passif** est également assujéti au risque de taux. En effet, l'engagement de l'entreprise d'assurance envers les assurés, par exemple, en assurance Vie est représentée par les **provisions mathématiques** qui correspondent à la valeur **actualisée** des engagements. En cas de choc de taux d'intérêt, la valeur économique du passif est impactée.

---

57. **Asset Liabilities Management** - ALM.

58. Les nouveaux contrats d'assurance-Vie sur le marché ne sont souvent plus assortis que d'une garantie en capital.

59. Il s'agit d'options implicites.

60. Garanties ex ante.



Dans un premier temps, l'entreprise d'assurance peut chercher à gérer son risque de taux d'intérêt via les techniques standards d'adossement<sup>61</sup> des flux ou via les techniques de couverture par sensibilité (duration-convexité).

Or le fait que

- les engagements peuvent être souvent plus **long** que la profondeur du marché d'actif.
- les contrats contiennent implicitement des options **implicites** plus ou complexes du fait de leur imbrication.

rendent les outils traditionnels non-implémentables ou qui donnent des valeurs sur/sous-estimées.

La construction d'un GSE sous probabilité historique et risque-neutre devient indispensable<sup>62</sup> pour pouvoir, par exemple, calculer la duration **effective** du passif ou pour évaluer **économiquement**<sup>63</sup> les options implicites.

La couverture du risque de taux d'intérêt **complète** sur l'ensemble des états du monde par l'entreprise d'assurance, serait extrêmement **coûteuse** et la charge serait difficilement acceptable pour l'assuré. Néanmoins, les entreprises d'assurance peuvent mettre en place une couverture **dynamique** avec la mise en place de couverture *au fil de l'eau* (*swaps, swaptions, ...*) en fonction de l'évolution des facteurs de risque. Le GSE permettra de tester l'**efficacité** des règles de couvertures mises en place sur l'horizon de projection.

### 1.2.3.2 Risque de crédit

Le risque de **défaut**<sup>64</sup> est le risque de **non**-paiement d'une dette (coupon et nominal d'une obligation). C'est un risque qui est devenu un risque prépondérant (crise de la dette, ...) en particulier pour les entreprises d'assurance qui ont une **allocation** principalement obligataire.

On définit également du risque de *spread* qui mesure l'**intensité** de la perception du risque de défaut et de liquidité d'un émetteur par les investisseurs. Ce risque peut être également mesuré par l'intermédiaire de notation<sup>65</sup>.

De la même façon que les taux d'intérêt, le GSE devient indispensable pour la prise en compte du risque de crédit dans les chroniques obligataires du modèle ALM.

### 1.2.3.3 Risque de rachats

Le **comportement** des assurés en assurance-Vie peut **dépendre** de l'évolution des facteurs de risques financiers. En effet, en fonction des **niveaux** de taux d'intérêt de marché, en particulier, et en **comparaison** avec les **taux servis** et ceux de la concurrence (fonds euros des autres entreprises d'assurance, livrets A bancaires, ...), des **rachats**<sup>66</sup> peuvent survenir.

---

61. Cash-flow matching.

62. Par exemple, l'estimation erronée des provisions mathématiques a conduit Equitable Life (GB) à la faillite 2001. *"Tout a commencé avec cette décision de tailler dans les bonis des polices de rentes à taux garanti de 90 000 sociétaires."*

63. Obligation de Solvabilité II.

64. Également appelé risque de signature.

65. **Rating**.

66. Rachats **conjuncturels** ou **dynamiques**. Les rachats **structurels** ne dépendent pas des facteurs financiers.

En environnement taux d'intérêt **bas** comme aujourd'hui<sup>67</sup>, un scénario de hausse de taux d'intérêt est très probable, les assurés (épargnants) seront dans ce cas susceptibles de racheter partiellement ou totalement leurs contrats pour profiter des nouveaux taux d'intérêt haussiers. Ainsi si des achats **massifs** se réalisaient, l'entreprise d'assurance devra **retourner** des positions au niveau de l'actif et pourrait réaliser des **moins-values** latentes ce qui affaiblirait sa solvabilité.

Afin de limiter ce risque de rachat les entreprises d'assurance peuvent<sup>68</sup> via les mécanismes de provisions pour participation aux bénéficiaires et de réserve de capitalisation **lisser** les taux servis dans le temps c'est-à-dire mener une politique stratégique afin de maintenir un rendement stable en fonction des cycles économiques.

Les GSE en monde **réel** permettront aux entreprises d'assurance un pilotage prospectif de leur risque de rachat en testant leurs règles de management (taux servis, ...), leur loi de rachats dynamiques. Les scénarios en **risque-neutre** permette d'évaluer l'option **implicite** de rachat dans le cadre de l'évaluation Best Estimate.

#### 1.2.3.4 Risque de liquidité

Le risque de **liquidité** mesure le fait qu'un **actif** puisse être acheté ou vendu **rapidement sans effet** significatif sur l'établissement de son **prix**. Ainsi un **débiteur** pourrait ne pas pouvoir rembourser sa dette, si la vente de l'actif adossé à son passif **ne** peut se réaliser. C'est une mesure **différente** du risque de **crédit** même si elle mesure, **également**, la capacité d'un débiteur à rembourser sa dette.

Ainsi, pour les actifs moins liquides, une **prime de liquidité** est donc **ajoutée** au prix pour compenser le risque pris par l'investisseur. Pour les titres obligataires, il s'agit d'un rendement **supplémentaire** au taux d'intérêt sans risque et au *spread* de crédit.

Pour les entreprises d'assurance, le risque de liquidité est un risque important à piloter. Par exemple, en assurance-Vie, les rachats **structurels**<sup>69</sup> ou **conjoncturels**<sup>70</sup> peuvent amener l'entreprise d'assurance à **devoir vendre** un actif pour ajuster son adossement actif-passif. Or, il peut être confronté, du fait de l'environnement économique et aux caractéristiques de la contrepartie, à la difficulté de vendre cet actif et de pouvoir rembourser ses assurés.

Pour mesurer ce risque de liquidité, les entreprises d'assurance peuvent établir un *gap*<sup>71</sup> de liquidité sur l'horizon de projection pour toute simulation monde-réel du GSE.

#### 1.2.3.5 Risque d'actifs de diversification

Les entreprises d'assurance doivent pouvoir quantifier l'impact des variations des actifs de diversifications (titres R332-20) :

— action

---

67. Cf. le document de l'ACPR [2].

68. En cas de fluctuations des marchés financiers modérés.

69. Il s'agit de rachat indépendants de la conjoncture économique.

70. ou dynamique.

71. Ecart entre les flux de l'actif et les flux du passif.

- entreprise de placement collectif<sup>72</sup>
- immobilier
- ...

En cas de scénario défavorable des actifs de diversification, l'Actif (du bilan) n'est plus adossé au Passif.

### 1.2.3.6 Allocation stratégique

L'**allocation stratégique** consiste à l'entreprise d'assurance de **déterminer** sa politique d'**investissement** permettant de réaliser un bénéfice tout en prenant en compte les contraintes de son passif (tarification, ...), de ses risques (risque de taux, ...) et des contraintes réglementaires (SCR, ...).

La détermination de l'allocation stratégique est devenu **déterminant** dans le cadre de la nouvelle directive Solvabilité II qui a introduit le risque **marché** dans le cadre de la valorisation **économique**. Ainsi le Capital Économique varie en fonction de l'évolution des postes de l'actif et donc **dépend** de l'allocation décidée par l'entreprise d'assurance.

Le développement de scénarios stochastiques des actifs permet de mesurer l'**impact** sur le modèle **ALM** de l'allocation (obligations et actifs de diversification) de l'entreprise d'assurance.

Du fait des **interactions** actifs-passifs, l'allocation stratégique nécessite une stratégie **multi-périodes**. En effet, à la fin de chaque période, la **composition** du bilan de l'entreprise d'assurance doit être **revue**. Les impacts sur le modèle ALM dépendent donc de la **réalisation**<sup>73</sup> des facteurs de risques dans la période précédente.

Les GSE **monde réel** deviennent un outil indispensable pour l'entreprise d'assurance afin de déterminer sa politique d'investissement et de couverture et de pilotage de ses ratios réglementaires.

### 1.2.3.7 Interaction actif-passif

Dans ce paragraphe nous allons voir comment l'évolution des grandeurs économiques conditionnant le prix des actifs peut impliquer des désinvestissements ou des investissements dus aux interactions actif-passif.

Nous allons introduire quelques définitions qui jouent un rôle important dans les **interactions actif-passif** en assurance-vie.

Par exemple, les contrats d'assurance-vie mono-support ou multi-supports (pour la partie fonds Euro) disposent de garanties favorables aux assurés : **Taux Minimum Garanti**<sup>74</sup> et la **Participation aux Bénéfices**<sup>75</sup> :

Le **taux d'intérêt technique** ou **taux minimum garanti**, représente la revalorisation **minimale** des provisions mathématiques que l'entreprise d'assurance garantit **chaque** année à l'assuré. Il est défini par le Code

---

72. Les OPC regroupent les OPCVM et les fonds d'investissements alternatifs (FCP)

73. Trajectoire sous probabilité historique

74. TMG = Taux Minimum Garanti.

75. PB = Participation aux Bénéfices.

des Assurances. Il peut être fixe ou variable et est **fixé** à la souscription du contrat.

Pour le fonds Euro, il s'agit du taux de rendement **minimum** servi à l'assuré chaque année.

La **participation aux bénéfices** correspond à la distribution aux souscripteurs d'assurance Vie des **bénéfices** réalisés grâce à leur épargne. Pour cette participation aux bénéfices, la réglementation<sup>76</sup> impose aux assureurs d'**attribuer** aux souscripteurs **85 % des bénéfices** financiers effectivement réalisés.

Néanmoins l'entreprise d'assurance peut choisir de **distribuer** une partie de la PB à la **fin de l'exercice** et/ou **doter** une partie en **Provision pour Participation aux Excédents**<sup>77\*</sup>. Cette provision, due aux assurés, devra être **entièrement** distribuée au bout de **huit** ans.

La PB dépend donc

- des **produits** financiers réalisés pendant l'exercice
- de la réglementation des 85% du Code des Assurances
- des règles de distributions **contractuelles**
- des règles de distribution commerciales (TMG boosté)
- de la **marge** de l'entreprise d'assurance

Le **taux servi** correspond à la partie de la PB qui est intégrée dans les PM.

La participation aux bénéfices **cible** est définie en fonction de la politique commerciale (nouveaux entrants, fidélisation de la clientèle). Le but est la **convergence** de la PB vers PB cible.

La **provision pour participation aux excédents** correspond à la partie des bénéfices qui **ne sont pas** distribuées **immédiatement** à l'assuré.

#### **Proposition 1.2.1 (Ventilation de la PB)**

*En fonction de l'écart entre la PB cible et la PB, la gestion actif-passif de l'entreprise d'assurance applique l'algorithme suivant :*

1. *Si la PB est **supérieure** à la PB cible alors*

- *la PB est **distribuée** à l'assuré c'est-à-dire intégrée dans les **provisions mathématiques***<sup>78</sup>
- *et l'excédent est **doté** à la PPE.*

---

76. Cf. Code des Assurances articles L. 331-3, A. 331-4.

77. PPE = Provision pour Participation aux Excédents.

78. PM = Provisions mathématiques

2. Si la PB est **inférieure** la PB cible c'est-à-dire que la production financière n'est pas suffisante afin d'atteindre le niveau de revalorisation cible, PB cible alors l'entreprise d'assurance pour pouvoir verser la PB cible peut

— réaliser des plus-values latentes des placements non-amortissables R 322-20.

— reprendre une partie de la PPE.

Si cela est possible l'entreprise d'assurance versera

— la PB cible est versé aux PM.

Dans le cas **contraire**, l'entreprise d'assurance versera

— la PB aux PM avec un *minimum* le TMG.

La formulation de la VIF (1.6) nécessite un modèle ALM intégrant les interactions actifs-passifs. Cf. le paragraphe 1.2.3.7 page 43.

En conclusion :

En conclusion, dans une approche prospective, le GSE sous probabilité historique permettra de distinguer les trajectoires **favorables** des actifs qui conduisent à une rendement **supplémentaire** (au TMG) pour les assurés.

## 1.2.4 ORSA

L'ORSA<sup>79</sup> est le sigle anglais pour l'évaluation interne des risques et de solvabilité.

Chaque entreprise d'assurance doit procéder à l'établissement d'un ORSA. Cette obligation est mentionnée dès le considérant 36 de la directive [69] Solvabilité II : "*chaque entreprise d'assurance et de réassurance devrait procéder régulièrement à l'évaluation de son besoin global de solvabilité, en tant que partie intégrante de sa stratégie commerciale et compte tenu de son profil de risque spécifique (évaluation interne des risques et de la solvabilité).*".

Dans ce cadre, l'entreprise d'assurance doit donc établir<sup>80</sup> en complément du pilier I, un processus interne<sup>81</sup> (**pilier II**) permettant

- **d'évaluer son besoin global de solvabilité** en prenant en compte son profil de risque **spécifique**. Les résultats de cette évaluation devront être un élément essentiel pour fixer sa politique stratégique afin de *planifier* les besoins futurs de solvabilité.

---

79. Own Risk Solvency Assesment.

80. Cf. l'article 45 de [69] Évaluation interne des risques et de la solvabilité.

81. Il existe également une dimension qualitative que nous présenterons pas.

- de **vérifier** si ses exigences en capital (SCR, MCR <sup>82</sup>) sont respectées de façon **permanente**. Les entreprises d'assurance doivent donc s'assurer qu'à chaque instant elles couvrent leur SCR.
- de **mesurer** comment son profil de risque s'écarte des hypothèses qui sous-tendent le calcul du capital de solvabilité requis dans le cadre du pilier 1.

Ce processus peut amener l'entreprise d'assurance à

- exiger un niveau de confiance **plus élevé** que celui du **pilier 1** (niveau de fonds propres à détenir afin d'éviter la ruine à un an dans 99,5% des cas.)
- considérer un horizon temporel **supérieur** à un an. Pluri-annuel, par exemple celui du plan stratégique moyen terme (PMT <sup>83</sup>).
- considérer d'**autres** risques (risque de défaut sur les obligations souveraines, risque de réputation, ...)
- utiliser une méthodologie différente de celle du pilier I en matière d'**agrégation** des risques.
- définir une stratégie d'**appétence** aux risques afin d'atteindre l'objectif de rentabilité définie par l'actionnaire.

Ainsi ce mécanisme d'ORSA nécessite un outil **prospectif** de génération de scénarios économiques.

Dans le cadre d'un exercice **prospectif pluri-annuel**, l'entreprise d'assurance peut être amené à calculer des  $SCR_u$  pour  $u = t + 1, u = t + 2, u = t + 3, \dots$ . Cet exercice entre dans le cadre d'un SdS dont l'horizon des simulations primaires n'est plus d'un an comme dans le calcul réglementaire du pilier I mais de deux ans, trois ans, ...

Le Pilier I exige au moins un calcul annuel <sup>84</sup>. Le respect **permanent** des exigences de capital quant à lui nécessite un calcul du SCR **plusieurs fois par an**. Néanmoins, du point de vue **opérationnel**, il est difficile de mettre en place un calcul complet infra-annuel du taux de couverture Solvabilité II dû au temps de calcul pour

- la **génération** des tables de scénarios économiques (calibration, ...)
- l'**intégration** dans le modèle ALM (model points, ...)

Ainsi les méthodes d'accélération peuvent être utilisées pour permettre un calcul du ratio de solvabilité sans effectuer un calcul complet.

---

82. Minimum Capital Requirement.

83. En général l'horizon du PMT est de 3 à 6 ans

84. Cf. l'article 102 de [69] Fréquence du calcul

## 1.3 Mise en oeuvre d'un GSE

Les paragraphes précédents ont montré la nécessité pour l'entreprise d'assurance de se doter d'un Générateur de Scénarios Économiques. Sa construction s'articule autour de plusieurs étapes communes mais dont les choix dépendent fortement de son usage qui en est fait.

1. **identifier** les facteurs de risque
2. **porter attention** aux choix des inputs (représentativité des données, fréquence des données).
3. **choisir** le modèle de dynamique pour chacun des facteurs de risque identifié
4. **sélectionner** la probabilité de simulation adéquate au contexte d'application du GSE.
5. **sélectionner** la structure de dépendance entre les trajectoires simulées.
6. **sélectionner** les méthodologies d'estimation et/ou de calibration.
7. **valider** les scénarios générés

### 1.3.1 Identification des facteurs de risques

La première étape de la construction d'un GSE consiste à identifier les sources (facteurs) de risques qui influent sur le résultat de l'entreprise d'assurance. Ce résultat est intimement liée aux caractéristiques du bilan et à la politique de gestion actif-passif inhérente.

Les facteurs de risques qui influent sur la valorisation de l'actif du bilan et via les règles de participations aux bénéfices (interaction actif-passif), la valorisation du passif sont ceux qu'on trouve généralement dans les GSE du fait de leur usage pour le calcul du SCR marché, forte composante du SCR global.

L'actif des bilan des entreprises d'assurance est principalement composé de

- titres obligataires souverains ou privées
- actions ou parts d'OPC
- investissements immobiliers

Ces facteurs sont donc principalement

- les **taux** d'intérêt également via l'actualisation

- le **crédit** (spread et défaut)
- les indices **actions** et **immobilier** et leur primes de risque inhérentes
- l'**inflation**, composante du taux d'intérêt nominal<sup>85</sup>

Le GSE devra donc modéliser les dynamiques de ces facteurs.

### 1.3.2 Inputs

Les caractéristiques des données pris en considération dépendent de la probabilité de simulation :

- Les inputs en **risque-neutre** sont principalement constitués de **prix** de produits financiers de calibration cotés sur les marchés financiers.
- Les inputs en **monde réel** sont constitués de données **historiques** des facteurs préalablement identifiés.

Mais quel que soit le type de simulation, le choix des données qui seront utilisées dans la construction du générateur de scénarios économiques est déterminant pour sa qualité :

- pour le risque-neutre : quels prix ? prix *mid*, prix de clôture ? quelle source ? Bloomberg , Reuters ? ...
- pour le monde réel : quelle plage historique ? quelle source ...

En particulier dans un cadre monde réel, une analyse **statistique** et **graphique** peut être opportun. Elle permet de donner une intuition sur la qualité de donnée, sur le type de modélisation qui pourrait être mise en place pour les facteurs choisis.

#### 1.3.2.1 Mesures statistiques usuelles

Les statistiques<sup>86</sup> descriptives souvent utilisées sont constitués des estimateurs empiriques des moments de la loi génératrice implicite :

- Les **minimum** et **maximum**<sup>87</sup>

$$x_{1:n} = \min(x_1, \dots, x_n) \quad \text{et} \quad x_{n:n} = \max(x_1, \dots, x_n)$$

- La **moyenne** empirique :

$$\hat{\mu} := n^{-1} \sum_{i=1}^n x_i$$

---

85. Cf. la relation de Fisher.

86. On suppose que l'échantillon historique est  $(x_1, \dots, x_n)$ .

87. Cf. l'annexe sur les statistiques d'ordre.



- La volatilité empirique<sup>88</sup> mesurant la dispersion autour de la moyenne :

$$\hat{\sigma} := \sqrt{(n-1)^{-1} \sum_{i=1}^n (x_i - \hat{\mu})^2}$$

- Les paramètres de forme :

- Le coefficient d'asymétrie de Pearson<sup>89</sup> empirique\* mesure l'asymétrie de la distribution empirique :

$$\hat{\gamma} := \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left( \frac{x_i - \hat{\mu}}{\hat{\sigma}} \right)^3$$

- Le coefficient d'aplatissement de Fisher<sup>90</sup> normalisé empirique\* traduit l'aplatissement de la distribution empirique par rapport à la distribution normale :

$$\kappa := \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \left( \frac{x_i - \hat{\mu}}{\hat{\sigma}} \right)^4 - 3 \frac{(n-1)^2}{(n-2)(n-3)}$$

### 1.3.2.2 Analyses graphiques - boîtes à moustaches

Il peut être intéressant de détecter dans les données les points **aberrants** qui apparaissent suite à une problématique de **qualité** de donnée. Néanmoins, ces points peuvent être réalisations d'**événements extrêmes**.

Voici ci-dessous un exemple<sup>91</sup> de boîtes à moustaches en assurance non-Vie sur les phénomènes des tempêtes sur la durée de tempêtes tropicales où la boîte à moustache permet d'illustrer le fait que le modèle de simulation a tendance à créer des durées cycloniques plus long.

---

88. On prend ici des estimateurs sans biais. idem pour les \*

89. **Skewness** - Par rapport à celui de **Pearson**, on utilise la normalisation par rapport à la valeur 3, valeur de référence de la loi normale.

90. **Kurtosis** - le coefficient d'aplatissement pour la loi normale est égal à 3.

91. Cf. le mémoire d'actuariat [64]

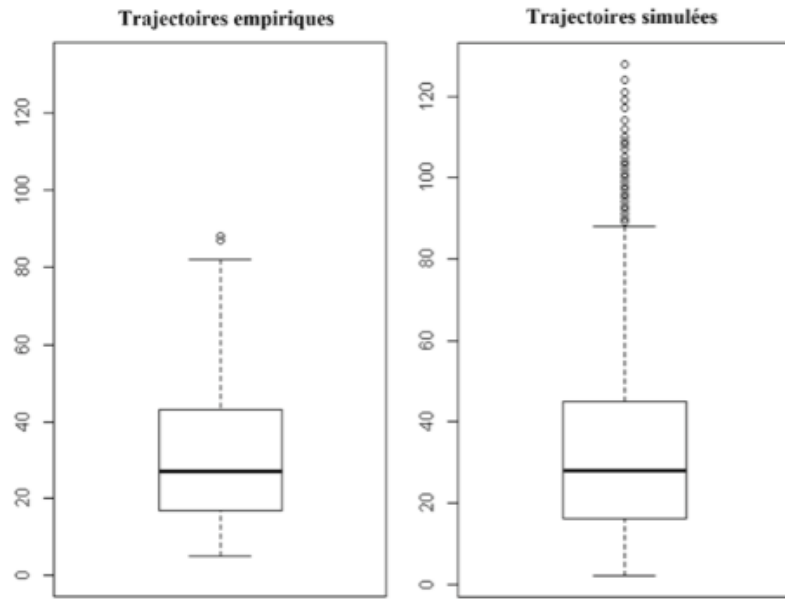


FIGURE 1.2 – Comparaison de données empiriques et simulées

### 1.3.2.3 Plage historique

Au delà de l'analyse descriptive, l'entreprise d'assurance doit porter une attention sur la **représentativité** de l'historique considéré.

Pour les simulations **monde-réel**, la **fréquence** et la **profondeur** historique des données d'estimation doivent être **cohérentes** avec l'**horizon** de **projection** envisagé.

Concernant la problématique de **fréquence**, pour les projections **long terme**, il convient de choisir plutôt une fréquence **faible** (annuelle par exemple voire mensuelle) une fréquence trop élevée (journalière par exemple) introduit trop de variabilité qui est plutôt en adéquation avec un pilotage de risque court terme.

Concernant la problématique de **profondeur historique**, la **cohérence**, entre les différentes **réalisations** de la même variable, **n'est pas** forcément évidente ce qui peut sous-tendre plusieurs **régimes** pour la même variable.

Par exemple, on peut remarquer dans le graphique suivant, un comportement de l'inflation annuelle radicalement différent sur la période récente et celle avant la première moitié du XX<sup>ème</sup> siècle :

Ainsi, dans ce cas, il ne semble pas cohérent sans retraitement d'utiliser une période aussi longue pour estimer les paramètres d'un modèle dynamique pour la variable inflation.

Le choix de la plage historique s'appuie donc souvent sur un avis d'**expert**. On peut essayer de corroborer

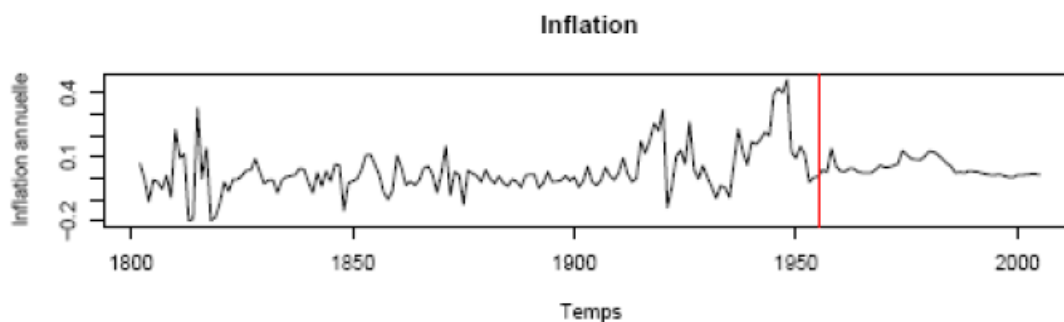


FIGURE 1.3 – Inflation annuelle sur la période 1800 - 2005<sup>92</sup>

cet avis via par exemple un test de **stabilité** des paramètres sur deux sous-population, par test de Chow<sup>93</sup>.

### 1.3.3 Modèles de projection

L'**horizon** de projection relativement **long** (en particulier en Vie) a des conséquences en terme de choix de modèle. Les modèles développés, en particulier en **monde-réel**, doivent être relativement cohérent avec la théorie économique (cycles, retour à la moyenne, signes de la variable ...).

De plus, l'entreprise d'assurance doit adapter la complexité de ses modèles, en particulier en **risque-neutre**, en fonction de la complexité de son portefeuille pour pouvoir par exemple valoriser les options implicites contenues dans ses contrats (par exemple les garantie de taux en assurance Vie).

D'autre part, comme nous l'avons vu dans les paragraphes précédents de calcul de mesure réglementaire liée à un quantile élevé, les modèles du générateur de scénarios économiques doivent pouvoir également générer des événements **extrêmes**.

### 1.3.4 Structure de dépendance

En assurance Vie et non-Vie, on observe des dépendances qui peuvent être liées à des éléments d'actifs, de passif ou entre l'actif et le passif<sup>94</sup>.

Ainsi au-delà du comportement marginal de chaque facteur de risque il est nécessaire d'étudier la dépendance entre les variables pour éviter de sous-estimer les risques.

Comme pour l'analyse descriptive pour les données univariées, il peut être opportun d'effectuer une analyse descriptive de la dépendance. Cette analyse est plutôt valable dans le cadre bi-dimensionnel.

93. Cf. la section sur les tests statistiques page 98.

93. Source : Cf. la présentation de F. Planchet [30]

94. Cette dernière dépendance, c'est ce qu'on a appelé interaction actif-passif

### 1.3.4.1 Analyse graphique : nuage de points

On peut commencer par tracer le **nuage de points** qui est défini par la représentation **graphique** sur le plan de l'échantillon historique  $(x_i, y_i)_{1 \leq i \leq n}$  décrivant la relation entre les deux variables  $X$  et  $Y$  :

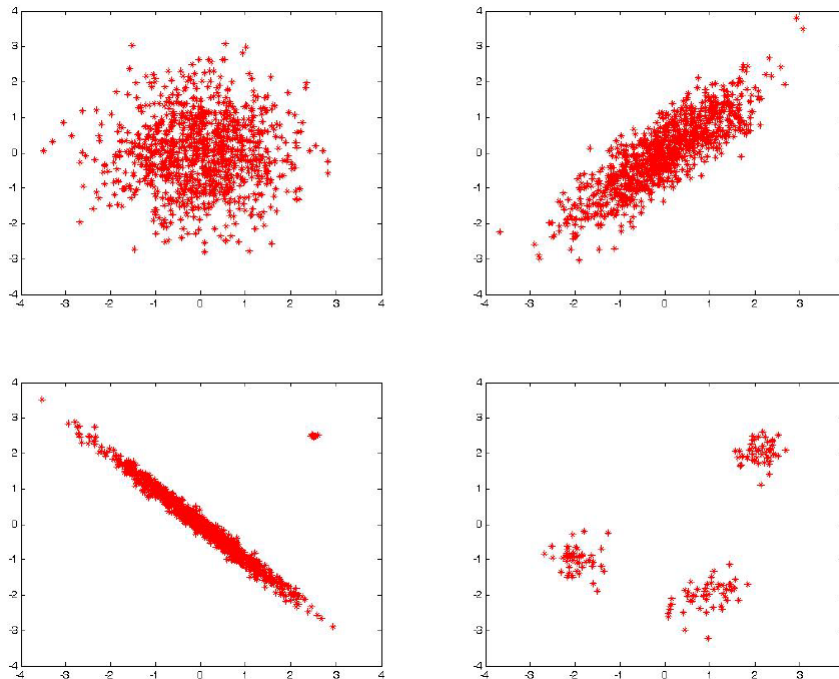


FIGURE 1.4 – Exemple de différents types de nuages<sup>95</sup> de points

La lecture graphique du nuage de point permet d'observer :

- les relations positives ou négatives
- la répartition plus ou moins homogène avec éventuellement des sous-groupes
- les données aberrantes

### 1.3.4.2 Mesures de dépendance : $\rho$ , $\rho_S$ , $\tau$

Comme pour l'analyse statistique univariée, il existe des statistiques mesurant la dépendance dans le cadre bi-dimensionnel.

---

94. Graphiques pris dans la présentation d'A. Popier Licence SPI.

La **corrélation linéaire**<sup>96</sup> **empirique** entre deux variables X et Y est donnée par :

$$\hat{\rho} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Souvent, pour les variables financières étudiées, cette mesure empirique n'est **pas** une bonne mesure de la **dépendance**.

En effet, les hypothèses **sous-jacentes** au calcul de la **corrélation linéaire** entre deux variables X et Y sont :

- la dépendance entre X et Y est **linéaire** donc une mesure **incapable** de capter des liaisons **non**-linéaires. Le contre-exemple usuel est la corrélation linéaire entre X et X<sup>2</sup> qui est nulle pour une variable de distribution gaussienne alors que la dépendance est manifeste.
- la variable (X,Y) suit une loi **normale** bi-variée. C'est donc une mesure de dépendance **paramétrique**.
- il y a **absence** d'auto-corrélation.

Or on observe dans la plus part des cas :

- Un rejet de l'hypothèse de normalité par le test de **Jarque-Bera**<sup>97</sup> par exemple. Les lois observées sont souvent à queues **épaisses**.
- Un **auto-corrélogramme** montrant des auto-corrélations.

Il existe d'autres mesure de dépendance **non-paramétrique**, le  $\rho_S$  de **Spearman** et le  $\tau$  de **Kendall**. Ce sont des mesures de **concordances**, c'est à dire qu'elle prennent des valeurs faibles, respectivement fortes **simultanément**. Les estimateur **empiriques** sont donnés<sup>98</sup> par :

$$\hat{\rho}_S := \frac{\sum_{i=1}^n (x_{(i:n)} - \bar{x})(y_{(i:n)} - \bar{y}_{(i:n)})}{\sqrt{\sum_{i=1}^n (x_{(i:n)} - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_{(i:n)} - \bar{y}_{(i:n)})^2}}$$

$$\hat{\tau} := \frac{2}{n(n-1)} \sum_{j=2}^n \sum_{i=1}^{j-1} \text{sign}((x_j - x_i)(y_j - y_i))$$

Ces coefficients, à la différence du coefficient de corrélation linéaire qui évaluent le lien entre les **valeurs** observées, évalue la corrélation entre les **rangs** des valeurs observées.

Le **tau** de Kendall entre deux variables aléatoires est égal à la **différence** entre la probabilité de concordance et la probabilité de non concordance.

Le **rhô** de Spearman entre deux variables aléatoires est égal à la corrélation linéaire entre les distributions **marginales** des variables.

96. Usuellement **corrélation** mais aussi corrélation de **Pearson** ou de **Bravais-Pearson**.

97. Cf. le paragraphe (B.2.0.3) page 151.

98.  $z_{(i:n)}$  représente la  $i^{\text{ème}}$  statistique d'ordre de la variable z et  $\bar{z}$  la moyenne empirique de la variable z.

### 1.3.4.3 Coefficients de dépendance de queues

Lorsqu'on s'intéresse à des mesures de risques d'événement extrêmes, on doit porter attention également aux dépendances extrêmes.

Les coefficients de dépendance de queue (à gauche et à droite) permettent de mettre en évidence des dépendances **extrêmes** entre deux variables.

On peut prendre l'exemple suivant en assurance non-Vie de dépendances des extrêmes avec une même corrélation :

Année de survenance	Charge sinistre évaluée		Année de survenance	Charge sinistre évaluée	
	Branche A	Branche B		Branche A	Branche B
1997	320	320	1997	320	825
1998	250	250	1998	250	850
1999	350	350	1999	350	850
2000	825	1507	2000	825	950
2001	850	1550	2001	850	320
2002	950	825	2002	950	250
2003	850	950	2003	850	350
2004	1550	850	2004	1550	1550
2005	1507	850	2005	1507	1507

• Estimation du coefficient de corrélation : 0,48

• Estimation du coefficient de corrélation : 0,48

FIGURE 1.5 – Exemple de dépendances des extrêmes<sup>99</sup>

Dans le tableau de gauche on observe une dépendance des valeurs faibles et à droite une dépendance des valeurs fortes tout en ayant la même corrélation.

Soit deux variables  $X$  et  $Y$  de distributions respectives  $F_X(\cdot)$  et  $F_Y(\cdot)$ . On définit ainsi la **dépendance de queue à droite**<sup>100\*</sup> pour un niveau de confiance  $\alpha$  comme la **probabilité** :

$$p^{\max} := \lim_{\alpha \rightarrow 1^-} \mathbb{P}\{X > F_X^{-1}(\alpha) \mid Y > F_Y^{-1}(\alpha)\}$$

C'est la probabilité que la variable  $X$  soit extrême **sachant** que la variable  $Y$  est également extrême. On dit que les variables  $X$  et  $Y$  sont **asymptotiquement dépendantes** à droite. Si cette probabilité est nulle, l'apparition d'un extrême pour la variable  $Y$  n'a **pas d'influence** sur l'apparition d'un extrême pour la variable  $X$ .

L'estimateur **empirique**<sup>101</sup> de cette probabilité est :

$$\hat{p}_\alpha^{\max} := \frac{\#\{x_i > q_X(\alpha) \cap y_i > q_Y(\alpha)\}}{\#\{y_i > q_Y(\alpha)\}}$$

99. Exemple de la présentation [13].

100. Supérieure - Upper tail dependency.

101.  $q_X(\alpha)$  et  $q_Y(\alpha)$  sont les quantiles empiriques.

On définit de la même manière la **dépendance de queue à gauche**<sup>102</sup> :

$$p^{\min} := \lim_{\alpha \rightarrow 0^+} \mathbb{P} \left\{ X \leq F_X^{-1}(\alpha) \mid Y \leq F_Y^{-1}(\alpha) \right\}$$

Nous venons de présenter des mesures de dépendances **bivariées**, dans le prochain paragraphe nous nous intéressons à la structure de dépendance de façon plus globale, la structure décrite par la loi **jointe**.

#### 1.3.4.4 Dépendance multivariée : copules

En assurance et en finance, de nombreux exemples comportent une composante **multivariée** :

- **en assurance Vie**, la tarification d'un contrat de rentes viagères sur une tête avec réversion fait l'hypothèse que les durées de vie des deux têtes sont indépendantes or on constate statistiquement une **sur-mortalité**<sup>103</sup> du conjoint au moment du décès de la première tête.
- **en assurance non-Vie**, la survenance d'un événements peut mettre en jeu les garanties de plusieurs contrats du même assuré. Par exemple, dans le cadre du risque de tempête, les montants des sinistres des contrats automobiles et multi-risques habitation sont liés.
- **en réassurance**, le risque de cumul peut être avéré si le le portefeuille d'un réassureur est composé principalement de contrats avec des cédantes pour le même risque.
- **en risque de crédit**, pour le *pricing* de produits dérivés de crédit<sup>104</sup> sur un panier de type *First-to-default*, CDO dépend de la loi jointe des temps de défaut des différents sous-jacents.
- **en risque de marché**, le calcul de la VaR d'un portefeuille dépend, par la somme, de la loi **jointe** des facteurs de risques :

$$\text{VaR} \left( \sum_{i=1}^d \omega_j S_{j,t} (\exp X_{j,t+1} - 1); 95\% \right)$$

où  $S_{j,t}$  est le  $j^{\text{ème}}$  facteur de risque et  $X_{j,t+1}$  est le log-rendement du  $j^{\text{ème}}$  facteur de risque.

- **en risque opérationnel**, la charge en capital au titre du risque opérationnel pour une banque CaR dépend de la loi **jointe** par la somme des pertes des  $J$  catégories de risque  $F(l_1, \dots, l_J)$  :

$$\text{CaR} := \text{VaR} \left( \sum_{j=1}^J l_j; 99,9\% \right)$$

où chaque  $l_j$  est une distribution Poisson-composée<sup>105</sup> déterminée par les lois de fréquence des pertes et de sévérité des pertes de chaque catégorie de risque.

---

102. Inférieure - Lower tail dependency

103. Syndrome du coeur brisé ou de tako-tsubo.

104. Credit derivatives.

105. Loss Distribution Approach :  $l_j = \sum_{k=1}^{N_j} \xi^k$

Dans une problématique multivariée, rappelons que la dépendance est **complètement** déterminée par la loi **jointe** mais il n'est souvent **pas** aisé d'estimer la loi **jointe**.

Le concept de **copule** introduite par A. Sklar<sup>106</sup> en 1959 permet d'exprimer une loi jointe plus **simple** :

$$\forall x \in \mathbb{R}^d \quad F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$$

La loi **jointe** est une **fonction** des lois **marginales**.

Les copules sont donc devenues un outil statistique très **populaire** lorsque l'on modélise des lois **multi-variées** en finance et en assurance.

On distingue les

- les copules **paramétriques**
- la copule **non-paramétrique**

Les copules **paramétriques** dépendent de un ou plusieurs **paramètres**. Principalement, elles se divisent en deux familles :

- copules **elliptiques**
- copules **archimédiennes**

Les copules **elliptiques** sont associées à une distribution **multivariées** dont la principale propriété est la **stabilité** par la **somme**<sup>107</sup>. On trouve comme exemples, les lois de **Gauss** et de **Student** multivariées.

Les copules **archimédiennes** sont associées aux distributions **multivariées** avec des **asymétries** de la dépendance de **queue** au **contraire** des copules elliptiques. On trouve comme exemples les copules de **Gumbel**<sup>108</sup>, de **Clayton**, de **Franck** ...

Après avoir présenté succinctement quelques copules paramétriques, nous vous présentons ici le concept de copule **empirique** définie à partir des **rangs** des observations du n-échantillon. Il s'agit donc d'une copule **non-paramétrique**. Cette copule est particulièrement utile dans le cadre de l'**inférence** statistique ou des tests d'**adéquation** des copules paramétriques : minimisation une distance<sup>109</sup> entre la copule empirique et la copule paramétrique. Mais dans notre mémoire nous allons plutôt l'utiliser comme **outil** (implicite) de **simulation**, ce sera l'objet du chapitre "Simulations multi-dimensionnelles" page 80.

Présentons maintenant une version empirique d'une loi multivariée définie par sa copule.

---

106. Cf. l'annexe pour une présentation des copules et pour une description détaillée Cf. le livre de Nelsen [51] de la théorie des copules ou le document du GRO [55].

107. La somme de variables de lois elliptiques est également une variable de loi elliptique.

108. Appelé également copule de **Gumbel-Hougaard**.

109. Exemple minimisation des statistiques de **Cramér-Von-Mises**, de **Kolmogorov-Smirnov** ou de **Cramér-Von-Mises** dans sa version  $L^1$ .



La distribution **multivariée empirique** du n-échantillon d-vectoriel  $(X_1, \dots, X_d)$  s'écrit

$$\widehat{F}_n(x_1, \dots, x_d) = n^{-1} \sum_{i=1}^n \prod_{j=1}^d 1_{\{x_j^i \leq x_j\}}$$

La distribution **univariée empirique** de la variable  $X_j$  est donnée par

$$\widehat{F}_{j,n}(x_j) = n^{-1} \sum_{i=1}^n 1_{\{x_j^i \leq x_j\}}$$

Or d'après la réciproque du théorème de Sklar (A.2) page 140 on a :

$$\forall u_1, \dots, u_d \in [0, 1]^d \quad C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$$

On peut donc écrire une approximation empirique<sup>110</sup> **naturelle** de la copule :

$$\begin{aligned} \forall u_1, \dots, u_d \in [0, 1]^d \quad \widehat{C}_n(u_1, \dots, u_d) &= \widehat{F}_n(\widehat{F}_{1,n}^{-1}(u_1), \dots, \widehat{F}_{j,n}^{-1}(u_d)) \\ &= n^{-1} \sum_{i=1}^n \prod_{j=1}^d 1_{\{x_j^i \leq \widehat{F}_{j,n}^{-1}(u_j)\}} \\ &= n^{-1} \sum_{i=1}^n \prod_{j=1}^d 1_{\{\widehat{F}_{j,n}(x_j^i) \leq u_j\}} \\ &= n^{-1} \sum_{i=1}^n \prod_{j=1}^d 1_{\{\frac{r_j^i}{n} \leq u_j\}} \end{aligned}$$

Les  $\{\frac{r_j^i}{n} : i = 1, \dots, n \text{ et } j = 1, \dots, d\}$  sont appelés **pseudo-observations**.

Du fait de la **discontinuité**<sup>111</sup> des fonctions de répartition empiriques, on a pas **unicité** de la copule.

En remarquant que cette copule empirique est **constante** sur le **treillis**<sup>112</sup> en dD (les pavés en 2D)

$$\left[\frac{i_1}{n}, \frac{i_1 + 1}{n}\right] \times \left[\frac{i_2}{n}, \frac{i_2 + 1}{n}\right] \times \dots \times \left[\frac{i_d}{n}, \frac{i_d + 1}{n}\right]$$

Paul Deheuvels propose de construire une copule<sup>113</sup> empirique en dD de la façon suivante :

### Définition 1.3.1 (Copule de Paul Deheuvels)

Soient

- $(x_1^i, \dots, x_d^i)_{1 \leq i \leq n}$  un échantillon de taille  $n$ , observation d'un  $n$ -uplet de variables aléatoires  $(X_1, \dots, X_d)$ .
- $(R_1^i, \dots, R_d^i)_{1 \leq i \leq n}$  la statistique de rang<sup>114</sup> associée à cet échantillon multivarié de dimension  $d$  i.e.

110. Où  $\widehat{F}_{j,n}^{-1}$  est le quantile empirique.

111. Cf. le théorème de Sklar, il y a **unicité** si les marginales sont continues.

112. On parle aussi d'**hypercube**.

113. Appelé par P. Deheuvels fonction de dépendance.

114. Rappelons que  $x_j^{(1:i)}$  est la  $i^{\text{ème}}$  statistique d'ordre de la variable  $X_j$ .

$(r_i^j)_{j=1, \dots, n}$  sont les rangs des  $(x_i^j)_{j=1, \dots, n}$ .

La copule  $\tilde{C}_n$  empirique de Paul Deheuvels est **définie** sur le treillis  $i_1, \dots, i_d \in \{1, \dots, n\}^d$  un **sous-ensemble** de  $[0, 1]^d$  :

$$\tilde{C}_n\left(\frac{i_1}{n}, \dots, \frac{i_d}{n}\right) = \frac{\text{nombre de } n\text{-uplets } (x_1, \dots, x_d) \text{ de l'échantillon tels que } x_1 \leq x_1^{(1:i_1)}, \dots, x_d \leq x_d^{(1:i_d)}}{n}$$

et plus généralement

$$\tilde{C}_n(u_1, \dots, u_d) = \begin{cases} \hat{C}_n\left(\frac{i_1}{n}, \dots, \frac{i_d}{n}\right) & \text{si } \frac{i_1}{n} \leq u < \frac{i_1+1}{n}, \dots, \frac{i_d}{n} \leq u < \frac{i_d+1}{n} \\ 1 & \text{si } i_1 = 1, \dots, i_d = 1 \end{cases} \quad (1.14)$$

C'est une loi **discrète**, en 2D nous avons la densité :

$$\tilde{c}_n\left(\frac{i}{n}, \frac{j}{n}\right) = \begin{cases} \frac{1}{n} & \text{si } (x_{(1:i)}, y_{(1:j)}) \in \{(x^1, y^1) \dots, (x^n, y^n)\} \\ 0 & \text{sinon} \end{cases}$$

Les copules  $C_n$  et  $\tilde{C}_n$  **coïncident** sur le treillis et la différence sur  $[0, 1]^d$  est au maximum

$$\sup_{u \in [0, 1]^d} |\hat{C}_n(u_1, \dots, u_n) - \tilde{C}_n(u_1, \dots, u_n)| = \frac{1}{n}$$

et on a la convergence du processus empirique

$$\mathbb{C}_n = \sqrt{n}(\tilde{C}_n - C)$$

### 1.3.4.5 Analyse graphique de dépendance de queues : dépendogramme

Le **dépendogramme** permet d'observer le caractère plus ou moins **simultané** des réalisations de l'échantillon. C'est une représentation graphique constituée du nuage de points des marginales empiriques extraites de l'échantillon de deux variables aléatoires. Or on a

$$\forall i = 1, \dots, n \quad \left(F_{1,n}(x_i), F_{2,n}(y_i)\right) \equiv \left(\frac{R_{1,n}(x_i)}{n}, \frac{R_{2,n}(y_i)}{n}\right)$$

Le dépendogramme est donc le **nuage de points** formé à partir des **statistiques de rang** issues de l'échantillon.

Rappelons que la copule empirique de Paul Deheuvels en 2D s'écrit

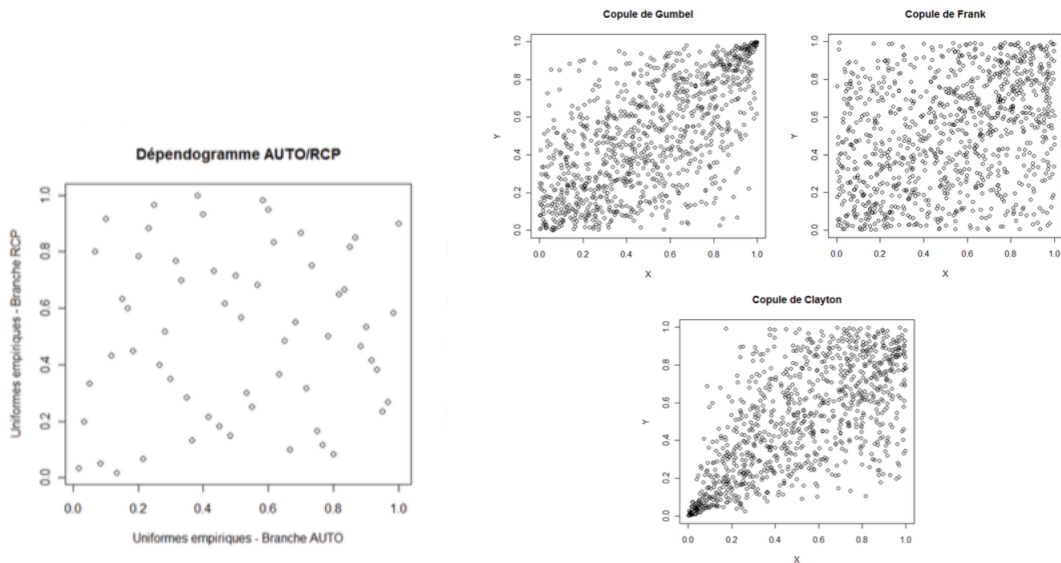
$$\forall k_1, k_2 \in 0, \dots, n \quad \hat{C}_n\left(\frac{k_1}{n}, \frac{k_2}{n}\right) = n^{-1} \sum_{i=1}^n 1_{\left\{\frac{r_1^i}{n} \leq \frac{k_1}{n}, \frac{r_2^i}{n} \leq \frac{k_2}{n}\right\}}$$

Les points du dépendogramme représente les valeurs prises par la copule de Paul Deheuvels.

On remarque que pour tout  $u_j \in [\frac{k_j}{n}, \frac{k_{j+1}}{n}]$  le nombre d'observations  $x_j^i$  tel que l'ensemble  $\{x_j^i < x_j^{(1:i)}\}$  soit non vide est identique.

On peut donc construire le dépendogramme pour deux variables X et Y à partir :

- soit d'un échantillon simulé à partir d'une variable aléatoire **théorique**
- soit d'un échantillon d'une variable **observée**



Le dépendogramme de gauche<sup>115</sup> représente la structure de dépendance observée entre deux branches de non-vie.

Le dépendogramme de droite représente la structure de dépendance issue de différentes copules **théorique** :

- Des points alignés sur l'**anti-diagonale**<sup>116</sup> traduit une dépendance **positive**.
- Des points alignés sur la **diagonale** traduit dépendance négative
- Copule de **Gumbel** : les points sont accentués en haut à gauche cela explicite la dépendance de **queue à droite**
- Copule de **Clayton** : les points sont accentués en bas à droite cela explicite une dépendance de **queue à gauche**
- Copule de **Franck** présente une dépendance **symétrique**.

115. Source Cf. [67].

116. Par identification au vocabulaire utilisé pour les matrices.

- Dans cet exemple de structure de dépendance entre deux branches non-vie semblerait plutôt être du type générée par la copule de **Franck**.

### 1.3.5 Probabilité de simulation

L'entreprise d'assurance devra bien choisir la bonne probabilité de simulation en fonction de la problématique soulevée <sup>117</sup>.

Mais les dynamiques choisies sous la probabilité historique et risque-neutre doivent être relativement **cohérentes**.

En effet, par exemple dans le cadre de processus **browniens** le théorème de **Girsanov** indique qu'entre deux probabilités équivalentes  $\mathbb{P}$  et  $\mathbb{Q}$ , il existe une densité de probabilité dite de **Radon-Nikodym** qui s'écrit comme une exponentielle <sup>118</sup> :

$$\frac{d\mathbb{Q}}{d\mathbb{P}}_{\mathcal{F}_t} = \exp\left\{-\int_t^T \lambda_u dW_u^{\mathbb{P}} - \frac{1}{2} \int_t^T \lambda_u^2 du\right\}$$

telle que  $W^{\mathbb{Q}} = dW^{\mathbb{P}} + \int_0^t \lambda_u du$  est **également** un mouvement brownien.

Dans le cadre d'un brownien géométrique la prime de risque instantané est égale au *spread* avec le taux de rendement sans risque par unité de volatilité :  $\frac{r-\mu}{\sigma}$

rajouter la condition de Novikov, et être moins affirmatif préciser le fait que l'on utilise des processus d'Ito et parler de la perte de l'unicité en cas d'incomplétude On ne peut donc pas choisir un modèle brownien avec saut sous la probabilité historique et un modèle brownien sans saut en risque-neutre.

Remarquons que la connaissance de la prime de risque permet de travailler dans les deux univers historique et risque-neutre.

### 1.3.6 Méthodologies d'estimation et/ou de calibration

En **monde-réel**, les **estimations** sont sensibles à la fenêtre de calibrage. Ainsi, les paramètres du modèle générateur devront être calibrés sur une **plage historique représentative** de l'état **actuel et futur probable** de la situation économique.

En **risque-neutre**, les calibrations marked-to-market dépendent du choix des instruments de calibrations.

### 1.3.7 Validation des scénarios générés

Dans le cadre de la mise en œuvre opérationnelle, les entreprises d'assurance ont tendance à vouloir **diminuer** le nombre de **simulations**. Or les résultats **théoriques** sont principalement **asymptotiques**.

Il est donc nécessaire d'étudier le nombre de simulations minimum pour assurer une robustesse des calculs basés sur les simulations.

---

117. Cf. les paragraphes précédents.

118. Exponentielle de Doléans-Dade.

Pour les GSE en monde **risque-neutre**, l'existence de la probabilité risque-neutre implique que la chronique (série temporelle) des prix actualisés est une **martingale**. On parle alors de **test martingale**. On peut également vérifier la **cohérence** des prix des produits **dérivés** avec les prix observés. On parle alors de **test de Market Consistency**. Cette validation des scénarios permet de prendre en compte le risque modèle.

Pour les GSE **monde réel**, il est plus **difficile** d'objectiver le **choix**. Néanmoins, par exemple dans le cadre d'un calcul de type SdS, on peut par exemple **augmenter** le nombre de scénarios monde réel **jusqu'à** *stabilisation* de l'estimation du Capital Économique.

## Chapitre 2

# Bootstrap naïf - un outil statistique

NOTRE GÉNÉRATEUR DE SCÉNARIOS ÉCONOMIQUES SERA CONSTRUIT À PARTIR DE **SIMULATIONS** BASÉES SUR LA MÉTHODE DU BOOTSTRAP.

L'OBJET DE CHAPITRE EST DE PRÉSENTER CET OUTIL STATISTIQUE INTRODUIT PAR **EFRON** :

- COMMENT CONSTRUIRE UN ESTIMATEUR BASÉ SUR LA DISTRIBUTION EMPIRIQUE ?
- COMMENT CALCULE-T-ON L'ERREUR-TYPE D'UN ESTIMATEUR PAR SIMULATION ?
- LA MÉTHODE DU BOOTSTRAP EST-ELLE CONVERGENTE <sup>1</sup> ?

---

1. Convergence de l'estimateur vers la vraie valeur.

Considérons une expérience du **hasard** engendrée par les variables aléatoires  $X_1, \dots, X_n$ . On suppose que ces variables  $X_i$  ont **même** loi qu'une variable dite **parent**  $X$ .

On observe une **réalisation** des valeurs prises par les variables aléatoires  $X_i$  que l'on note  $x_1, \dots, x_n$ . Ces réalisations représente un **sous-ensemble** d'une population **abstraite** constitué de toutes les observations possibles de la variable  $X$ .

On appelle **n-échantillon**  $\mathbb{X}$ , le **n-uplet**  $(X_1, \dots, X_n)$  où les variables aléatoires  $X_i$  sont **indépendantes** et toutes de **mêmes** lois, loi de la variable parent  $X$ .

On note  $x := (x_1, \dots, x_n)$  la **réalisation** de cet n-échantillon.

Dans beaucoup de problèmes **statistiques**, on est intéressé par tirer des **conclusions** à partir d'un nombre **limité** de données ( $n < +\infty$ ) constituées par la réalisation du n-échantillon  $\mathbb{X}$ .

On souhaite donc déterminer les caractéristiques d'une population **non-observée entièrement** à partir d'un échantillon de taille **finie**.

On suppose souvent que cette population peut être **décrite** à partir d'un **paramètre**<sup>2</sup>  $\theta$  c'est-à-dire que la connaissance de ce paramètre  $\theta$  permet la connaissance de l'ensemble de la population.

L'**inférence** statistique consiste à **déterminer** ce paramètre  $\theta$ .

On appelle **statistique**, la **fonction**  $t_n$  des valeurs réalisées  $x := (x_1, \dots, x_n)$  qui permette d'estimer  $\theta$  :

$$\begin{aligned} t_n : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x = (x_1, \dots, x_n) &\longrightarrow \hat{\theta}(x) := t_n(x_1, \dots, x_n) \end{aligned}$$

Si  $x$  est la réalisation de la variable aléatoire  $\mathbb{X}$  alors  $t_n(x)$  est la réalisation de la statistique  $T_n := t_n(\mathbb{X})$ .

On peut donner les exemples suivants :

- la moyenne empirique :  $t_n(x) := \bar{x}$
- la statistique d'ordre 1 :  $t_n(x) := x_{(1)}$
- le vecteur de  $\mathbb{R}^3$  :  $t_n(x) := (x_2, -\ln x_4, x_9^2)$

L'estimation consiste à répondre principalement aux problèmes suivants :

- Précision de l'estimateur
- biais du  $\hat{\theta}$

$$\mathbb{E}[t_n(\mathbb{X}) - \theta]$$

espérance **sous** la probabilité déterminée par la distribution  $F_\theta$  fonction de répartition des  $X_i$ .

---

2. Population parameter.

— variance de  $\widehat{\theta}$

$$\text{Var}[t_n(\mathbb{X})]$$

variance **sous** la probabilité déterminée par la distribution  $F_\theta$  fonction de répartition des  $X_i$ .

— Test d'hypothèse sur le paramètre  $\theta$

## 2.1 Méthode d'estimation par injection : Estimateurs *plug-in*

Soit  $F_\theta$  la fonction de répartition de la loi de l'échantillon  $(X_1, \dots, X_n)$  i.e. loi de la variable parent  $X$ .

On appelle **fonctionnelle statistique**<sup>3</sup>, tout paramètre s'exprimant comme une fonctionnelle  $\Psi$  de la distribution  $F_\theta$  :

$$\theta = \Psi(F_\theta)$$

On peut considérer les exemples :

— la moyenne  $\mu = \int x dF_\theta(x) =: \Psi_\mu(F_\theta)$

— la variance  $\sigma^2 = \int (x - \mu)^2 dF_\theta(x) =: \Psi_\sigma(F_\theta)$

— le quantile  $\alpha_p = F^{-1}(p) =: \Psi_{\alpha_p}(F_\theta)$

— ...

Dans ce cadre, on peut obtenir un estimateur de  $\theta$  via une méthode **non-paramétrique** à partir du n-échantillon.

Introduisons la notion de mesure **empirique**.

Si on considère que les variables aléatoires  $X_1, \dots, X_n$  sont **identiquement** et **indépendamment** distribuées à partir d'une distribution (fonction de répartition<sup>4</sup>) **inconnue**, on définit la fonction de distribution **empirique**<sup>5</sup> comme la **proportion**<sup>6</sup> issue du n-échantillon :

$$\widehat{F}_n : \mathbb{R} \longrightarrow [0, 1]$$

$$x \longrightarrow \frac{\#\{x_i \leq x\}}{n} = n^{-1} \sum_{i=1}^n 1_{x_i \leq x} = n^{-1} \sum_{i=1}^n H(x - x_i)$$

La fonction de répartition empirique attribue une probabilité **identique** à chaque élément  $x_i$  de l'échantillon.

---

3. Statistical functional.

4. Cumulative distribution function CDF.

5. Empirical distribution function EDF.

6. On note  $\#\{A\}$  le cardinal de l'ensemble  $A$ . On note  $H$  la fonction **Heaviside** définie par

$$H(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$



Cette fonction est en escalier, on observe des sauts de taille  $n^{-1}$  en chaque point de l'échantillon.

Le paramètre  $\theta$  peut être estimé par un estimateur **non**-paramétrique construit à partir de la fonction de répartition empirique, l'estimateur *plug-in*.

L'estimateur *plug-in* de  $\theta = \Psi(F)$  est défini par

$$\widehat{\theta}_n = \Psi(\widehat{F}_n) \quad (2.1)$$

Par exemple, l'estimateur *plug-in* de la moyenne  $\mu = \Psi(F)$  est donné<sup>7</sup> par :

$$\begin{aligned} \widehat{\mu}_n &= \Psi(\widehat{F}_n) \\ &= \int x d\widehat{F}_n(x) \\ &= n^{-1} \sum_{i=1}^n \int x dH(x - x_i) \\ &= n^{-1} \sum_{i=1}^n \int x \delta_{x_i} dx = n^{-1} \sum_{i=1}^n x_i \\ &= \bar{x} \end{aligned}$$

L'**erreur type**<sup>8</sup> de l'estimateur<sup>9</sup> de la moyenne i.e. l'écart-type de la statistique  $T_n := t_n(X_1, \dots, X_n) := \bar{X}$  dont la valeur réalisée est  $\widehat{\mu}_n$ , est défini par :

$$se_F(T_n) = \text{Var}(\bar{X}) = \frac{\sigma^2}{n}$$

Puisque l'estimateur *plug-in* de la variance est donné par

$$\begin{aligned} \widehat{\sigma}^2 = \Psi'(\widehat{F}_n) &= \int x^2 d\widehat{F}_n(x) - \left( \int x d\widehat{F}_n(x) \right)^2 \\ &= n^{-1} \sum_{i=1}^n x_i^2 - \bar{x}^2 = n^{-1} \sum_{i=1}^n (x_i - \bar{x})^2 \end{aligned}$$

Alors la mesure de l'erreur type est donné par :

$$se_{\widehat{F}_n}(T_n) = \widehat{se}_F(\bar{X}) = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n^2}} \quad (2.2)$$

On retrouve le résultat **usuel** de l'estimateur de variance empirique<sup>10</sup>  $\bar{X}$ .

7. On rappelle que  $\widehat{F}_n$  est la fonction de répartition empirique.  $\delta$  est la fonction de **Dirac** - the delta distribution.

8. Standard error.

9. On rappelle que les variables  $X_i$  sont **i.i.d.**

10. Estimateur avec biais.

## 2.2 Algorithme du Bootstrap naïf d'Efron

Dans ce paragraphe, nous allons nous focaliser sur la question des propriétés de la distribution de la statistique  $T$ . Par exemple, quel est le biais, l'erreur-type ou quels sont les quantiles ?

Dans la précédente section, nous avons calculé l'erreur-type (équation 2.2) de la statistique  $\bar{X}$  en appliquant directement la formule de Bienaymé pour la somme de variables aléatoires non corrélées et la méthode par injection.

Mais plus généralement on ne peut pas déterminer directement la variance :

$$se_F^2(T_n) = \int (t_n(x_1, \dots, x_n) - \mathbb{E}_F[T_n])^2 dF(x_1) \cdots dF(x_n)$$

où

$$\mathbb{E}_F[T_n] = \int t_n(x_1, \dots, x_n) dF(x_1) \cdots dF(x_n)$$

Même si l'on utilise l'estimateur *plug-in*, il peut être difficile de la calculer :

$$se_{\widehat{F}_n}^2(T_n) = \int (t_n(x_1, \dots, x_n) - \mathbb{E}_{\widehat{F}_n}[T_n])^2 dF(x_1) \cdots dF(x_n)$$

où

$$\mathbb{E}_{\widehat{F}_n}[T_n] = \int t_n(x_1, \dots, x_n) d\widehat{F}_n(x_1) \cdots d\widehat{F}_n(x_n)$$

Ainsi, nous allons introduire<sup>11</sup> la méthode de **Bootstrap** proposé par **Efron** comme outil statistique numérique pour **approximer** l'erreur type de la statistique avec un estimateur basé sur des **simulations**.

L'algorithme du *Bootstrap* ne requiert pas de calcul théorique pour calculer l'écart-type et demeure valide même si la statistique  $T_n := t_n(X_1, \dots, X_n)$  est **complexe**.

Les étapes\* de l'algorithme de *Bootstrap* sont les suivantes :

**Étape 1 :** CONSTRUIRE LA DISTRIBUTION empirique  $\widehat{F}_n$  À PARTIR DE L'**observation** D'UN ÉCHANTILLON EN ATTRIBUANT LA PROBABILITÉ **identique**  $n^{-1}$  À CHAQUE OBSERVATION,  $x_1, \dots, x_n$  DE L'ÉCHANTILLON.

$$\mathbb{P}\{X_j^* = x_i\} = n^{-1} \quad \text{pour tout } j \text{ et } i$$

**Étape 2 :** ENSUITE, À PARTIR DE LA FONCTION DE RÉPARTITION EMPIRIQUE  $\widehat{F}_n$ , TIRER (ALÉATOIREMENT) **avec remise** UN ÉCHANTILLON DE TAILLE  $n$ . C'EST L'ÉCHANTILLON *Bootstrap*, DONT LES RÉALISATIONS SONT NOTÉES  $x_1^*, \dots, x_n^*$ .

**Étape 3 :** ENFIN, **calculer** LA VALEUR DE LA STATISTIQUE  $t$  DE L'échantillonnage  $t_n^*$

$$t_n^* := t_n(x_1^*, \dots, x_n^*)$$

---

11. Pour une description plus détaillée, nous renvoyons le lecteur au livre d'**Efron** [46].

**Étape 4 : Répéter**  $B$  FOIS LES DEUX ÉTAPES PRÉCÉDENTES AFIN DE **construire** UN ÉCHANTILLON EMPIRIQUE (**simulé**) DE  $B$  VALEURS DE LA STATISTIQUE  $t_n^*$ . NOUS DÉFINISSONS AINSI LA DISTRIBUTION EMPIRIQUE DE LA STATISTIQUE  $t_n^*$  :

$$\mathbb{P}_{\widehat{F}_n} \{T_n^* \leq x\}$$

**Résultat :** ON OBTIENT **une** ESTIMATION, SIMULÉE, DE L'ERREUR-TYPE DE LA STATISTIQUE  $T_n$  :

$$se_B^2 = B^{-1} \sum_{i=1}^B t_n^*(x_1^*, \dots, x_n^*) - \bar{t}_n^*$$

## 2.3 Convergence du *Bootstrap*

Après avoir décrit l'algorithme du *Bootstrap*, nous introduisons maintenant les notations suivantes pour illustrer la **consistance** du *Bootstrap*. Dans beaucoup de problèmes statistiques nous sommes intéressés à estimer la distribution  $\widehat{\theta}_n$  sous la distribution  $F_\theta$  :

$$F_{\widehat{\theta}_n}(x) := \mathbb{P}_{F_\theta} \{ \sqrt{n}(\widehat{\theta}_n - \theta) \leq x \}$$

où <sup>12</sup>

$$\widehat{\theta}_n = \Psi(\widehat{F}_n) = t_n(X_1, \dots, X_n)$$

L'idée de la méthode du *Bootstrap* est d'**approximer** cette distribution par la distribution **conditionnelle** :

$$F_{\widehat{\theta}_n^*}(x) := \mathbb{P}_{\widehat{F}_n} \left\{ \sqrt{n}(\widehat{\theta}_n^* - \widehat{\theta}_n) \leq x \mid X_1 = x_1, \dots, X_n = x_n \right\}$$

Pour évaluer l'approximation, considérons la quantité définie par la **distance** de Kolmogorov :

$$K(F_{\widehat{\theta}_n}, F_{\widehat{\theta}_n^*}) := \| F_{\widehat{\theta}_n}(x) - F_{\widehat{\theta}_n^*}(x) \|_\infty$$

On dit que le *Bootstrap* est **consistant** si

$$\lim_{n \rightarrow +\infty} K(F_{\widehat{\theta}_n}, F_{\widehat{\theta}_n^*}) = 0$$

Considérons le cas <sup>13</sup> où  $\theta = \mu$  i.e. on s'intéresse à

$$R_n(\mathbb{X}, F_\theta) = \sqrt{n}(\widehat{\theta} - \theta)$$

Nous avons <sup>14</sup> :

$$\begin{cases} \mathbb{E}_F[\widehat{\theta}_n] = \mathbb{E}_F[\bar{X}] = \theta \\ \text{Var}_F(\widehat{\theta}_n) = \text{Var}_F(\bar{X}) = \frac{\sigma^2}{n} \end{cases}$$

12. Remarquons que  $\theta_n$  est une variable aléatoire.

13. On peut trouver une généralisation dans les articles donner les références.

14. On rappelle que les  $X_i$  sont i.i.d de  $F$  avec  $\mathbb{E}[X_i] = \theta$ .

et

$$\begin{cases} \mathbb{E}_{\widehat{F}_n}[\widehat{\theta}_n^*] = \mathbb{E}_{\widehat{F}_n}[\bar{X}^*] = n^{-1} \sum_{i=1}^n \mathbb{E}_{\widehat{F}_n}[X_i^*] = \mathbb{E}_{\widehat{F}_n}[X_1^*] = \bar{x} \\ \text{Var}_{\widehat{F}_n}(\widehat{\theta}_n^*) = \text{Var}_{\widehat{F}_n}(\bar{X}^*) = n^{-1} \text{Var}_{\widehat{F}_n}(X_1^*) = n^{-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{\widehat{\sigma^2}}{n} \end{cases}$$

Considérons les trois variables suivantes :

$$R_n := \sqrt{n}(\widehat{\theta}_n - \theta)$$

$$R_n^* := \sqrt{n}(\widehat{\theta}_n^* - \bar{x})$$

$$K_n := \sup_{x \in \text{Supp}(F)} |\mathbb{P}_F(R_n \leq x) - \mathbb{P}_{\widehat{F}_n}(R_n^* \leq x)|$$

Maintenant, nous devons montrer<sup>15</sup> que  $\lim_{n \rightarrow +\infty} K_n = 0$  en calculant la borne supérieure<sup>16</sup> :

$$\begin{aligned} K_n &\leq \sup_{x \in \text{Supp}(F)} |\mathbb{P}_F\{\frac{R_n}{\sigma} \leq \frac{x}{\sigma}\} - \Phi(\frac{x}{\sigma})| \\ &+ \sup_{x \in \text{Supp}(F)} |\Phi(\frac{x}{\sigma}) - \Phi(\frac{\widehat{x}}{\widehat{\sigma}})| \\ &+ \sup_{x \in \text{Supp}(F)} |\mathbb{P}_{\widehat{F}_n}\{\frac{R_n^*}{\widehat{\sigma}} \leq \frac{x}{\widehat{\sigma}}\} - \Phi(\frac{x}{\widehat{\sigma}})| \end{aligned}$$

On va maintenant s'intéresser au trois membres de droite de l'inégalité :

$$\boxed{K_n \leq A_n + B_n + C_n}$$

D'après la borne de Berry-Essen<sup>17</sup> du théorème centrale limit nous avons

$$A_n = \sup_{x \in \text{Supp}(F)} |\mathbb{P}_F\{\frac{R_n}{\sigma} \leq \frac{x}{\sigma}\} - \Phi(\frac{x}{\sigma})| \leq \frac{C \mathbb{E}_F[|X_1 - \theta|^3]}{\sigma^3 \sqrt{n}}$$

Cela montre clairement la convergence uniforme c'est-à-dire  $\lim_{n \rightarrow +\infty} A_n = 0$  si le moment d'ordre 3 existe et si les  $X_i$  sont **i.i.d.**.

L'estimateur *plug-in* de  $\sigma^2$  peut être écrit de la façon suivante  $n^{-1} \sum_{i=1}^n (X_i - \bar{X})^2 = n^{-1} \sum_{i=1}^n X_i^2 - \bar{X}^2$  et donc d'après la loi forte des grands nombres<sup>18</sup> (de Kolmogorov) nous avons :

$$\bar{X} \xrightarrow{a.s.} \theta \text{ and } n^{-1} \sum_{i=1}^n X_i^2 \xrightarrow{a.s.} \mathbb{E}_F[X_1^2]$$

15. d'après l'inégalité triangulaire et en introduisant la distribution normale centrée réduite  $\Phi(x)$

16. Upper bound.

17. Donner une référence.

18. *Strong law of large numbers* Cf. le livre de [49] page 73.

alors l'estimateur *plug-in*  $\widehat{\sigma}^2$  converge presque sûrement<sup>19</sup> vers  $\sigma^2$  et d'après le théorème continuous mapping<sup>20</sup> nous avons la convergence uniforme :

$$\lim_{n \rightarrow +\infty} B_n = \lim_{n \rightarrow +\infty} \sup_{x \in \text{Supp}(F)} \left| \Phi\left(\frac{x}{\widehat{\sigma}}\right) - \Phi\left(\frac{x}{\sigma}\right) \right| = 0$$

Enfin, de nouveau en utilisant la borne de Berry-Essen nous avons :

$$\begin{aligned} C_n &= \sup_{x \in \text{Supp}(F)} \left| \mathbb{P}_{\widehat{F}_n} \left\{ \frac{R_n^*}{\widehat{\sigma}} - \Phi\left(\frac{x}{\widehat{\sigma}}\right) \right\} \right| \leq \frac{C^* \mathbb{E}_{\widehat{F}_n} [|X_1^* - \theta|^3]}{\widehat{\sigma}^3 \sqrt{n}} \\ &= \frac{C^*}{\widehat{\sigma}^3 \sqrt{n}} \sum_{i=1}^n |x_i - \bar{x}|^3 \\ &\leq \frac{2^3 C^*}{\widehat{\sigma}^3} \left( n^{-3/2} \sum_{i=1}^n |x_i - \theta|^3 + n^{-1/2} |\theta - \bar{x}|^3 \right) \end{aligned}$$

Et nous terminons la démonstration<sup>21</sup> en appliquant la loi forte des grand nombres de Zygmund-Marcinkiewicz<sup>22</sup> :

$$n^{-3/2} \sum_{i=1}^n |X_i - \theta|^3 \xrightarrow{a.s.} 0 \text{ and } n^{-1/2} |\theta - \bar{x}|^3 \xrightarrow{a.s.} 0$$

---

19. Almost sure convergence a.s. .

20. Cf. le théorème 4.27 de Mann et al. dans le livre de [49] page 76.

21. On rappelle que la distribution  $F_{\widehat{\theta}_n^*}$  représente la distribution conditionnelle où l'on a remplacé respectivement les  $x_i$  par  $X_i$ .

22. Cf. le théorème 4.23 du livre de [49] page 73.

## Chapitre 3

# Simulations univariées

L'OBJET DE CE CHAPITRE EST DE PRÉSENTER NOTRE CONSTRUCTION DE SCÉNARIOS POUR UN **SEUL** ACTIF :

- COMMENT UTILISE-T-ON LE BOOTSTRAP COMME OUTIL DE **SIMULATION** ?
- COMMENT INTRODUIRE DES ÉVÈNEMENTS **EXTRÊMES** DANS NOS SIMULATIONS ?

NOUS OBTIENDRONS UN ALGORITHME **RÉCURSIF** GÉNÉRANT À CHAQUE PAS DES SCÉNARIOS DONT LA DISTRIBUTION EST UNE LOI **MÉLANGE**.

### 3.1 Bootstrap de résidus d'un modèle économétrique

Un scénario **historique** d'une variable économique peut être modélisé par une **série temporelle**.

Pour créer de **nouveaux** scénarios nous souhaitons utiliser la faculté *Bootstrap* naïf de **créer de nouveaux** échantillons à partir des valeurs historiques que constitue la série temporelle.

Le *Bootstrap* naïf d'Efron n'est **valide**<sup>1</sup> que pour des variables aléatoires **identiquement et indépendamment distribuée**. Or les séries temporelles présentent souvent le caractère d'**auto-corrélation**. Nous ne pouvons donc pas *bootstraper* les valeurs historique de la série temporelle.

Nous **contournerons** cette hypothèse nécessaire en *bootstrapant* les **résidus** observés d'un modèle **économétrique** modélisant la série temporelle historique. En effet, si le modèle économétrique est de *bonne qualité*, les résidus de ce modèle auront le caractère i.i.d.

Voici notre algorithme basé sur le *Bootstrap* permettant de créer une nouvelle série temporelle :

**étape 1** Sélectionner un modèle économétrique.

**étape 2** Estimer le modèle pour obtenir les coefficients estimés.

**étape 3** Calculer les résidus centrés

**étape 4** Effectuer un tirage aléatoire avec une probabilité  $n^{-1}$  un résidu observé.

**étape 5** Injecter cette valeur dans le modèle économétrique pour déterminer une nouvelle valeur de la série temporelle.

**étape 6** Répéter B fois l'étape 4 pour obtenir une distribution empirique de la valeur consécutive de la série temporelle  $X_{t+1}^*(\omega)$

**étape 7** Répéter n fois l'étape 4 et 5 pour obtenir une nouvelle trajectoire (scénario)  
 $X_{t+1}^*(\omega_1), \dots, X_{t+n}^*(\omega_n)$

---

1. Cf. Chapitre 2.

Pour illustrer cet algorithme, prenons par exemple, un échantillon historique de résidus centrés estimés via un processus AR(1) :  $X_{t+1} = X_t + \epsilon$

$$(\epsilon_1, \dots, \epsilon_H)$$

A partir de cet échantillon nous pouvons créer  $n$  échantillon bootstrap constitué d'un ré-échantillonnage avec remise équi-probable de l'échantillon ci-dessus :

$$\begin{aligned} & (\epsilon_1^*(\omega_1), \dots, \epsilon_H^*(\omega_1)) \\ & \quad \vdots \\ & (\epsilon_1^*(\omega_n), \dots, \epsilon_H^*(\omega_n)) \end{aligned}$$

Ensuite, on utilise la procédure **réursive** induite par le modèle  $AR(1)$  permettant d'obtenir de nouvelles valeurs de la série temporelle :

$$\left\{ \begin{array}{l} X_{t+1}^* = \hat{\alpha} + \hat{\phi}X_t + \epsilon_i^*(\omega_1) \\ X_{t+2}^* = \hat{\alpha} + \hat{\phi}X_{t+1}^* + \epsilon_j^*(\omega_2) \\ \text{et ainsi de suite} \\ X_{t+n}^* = \hat{\alpha} + \hat{\phi}X_{t+n-1}^* + \epsilon_k^*(\omega_n) \end{array} \right.$$

où les  $\epsilon_i^*$ ,  $\epsilon_j^*$  et  $\epsilon_k^*$  sont tirés **indépendamment** de leur échantillon **respectif**.



## 3.2 Modification des queues de distributions de la loi Bootstrap

### 3.2.1 Loïs des extremums généralisés - GEV

Nous venons de voir dans la section précédente que la méthode du *Bootstrap* permettait de simuler des trajectoires basées sur l'**historique**, cependant nous **souhaitons** également générer des scénarios extrêmes<sup>2</sup> **non encore observés** tout en étant cohérent avec l'historique. Or la méthode du *Bootstrap naïf* échoue lorsque

— le moment d'ordre 2 n'est pas fini i.e.  $\mathbb{E}[X_i^2] = +\infty$

— pour les statistiques extrêmes

$$T_n(X_1, \dots, X_n) := \max_{0 \leq i \leq n} X_i$$

Pour cela, nous allons utiliser la théorie des valeurs extrêmes<sup>3</sup> dans notre générateur *bootstrap* pour **pouvoir** intégrer des scénarios **extrêmes** dans les trajectoires *Bootstrap* sous la probabilité historique.

Nous **introduisons** succinctement dans ce paragraphe la théorie de valeurs extrêmes univariées mais pour une présentation complète, nous renvoyons le lecteur à l'ouvrage de **référence** de Embrechts et al. [47].

La distribution<sup>4</sup> de la  $n^{\text{ème}}$  statistique d'ordre  $X_{(n)}$  est donnée par  $F_X^n$ , en effet :

$$\begin{aligned} \mathbb{P}\{X_{(n)} \leq x\} &= \mathbb{P}\{X_1 \leq x, \dots, X_n \leq x\} \\ &= F_X(x)^n \end{aligned}$$

ceci implique que la distribution limite de  $X_{(n)}$  est **dégénérée** :

$$\lim_{n \rightarrow +\infty} \mathbb{P}\{X_{(n)} \leq x\} = \begin{cases} 1 & \text{si } F_X(x) = 1 \\ 0 & \text{si } F_X(x) < 1 \end{cases}$$

Le **théorème central limite** indique que la distribution de la **moyenne empirique normalisée**<sup>5</sup> tend vers une loi non-dégénérée :

$$\frac{n^{-1} \sum_{i=1}^n X_i - \mathbb{E}[X_i]}{n^{-\frac{1}{2}} \sqrt{\text{Var}(X_i)}} \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}(0, 1)$$

De la même façon mais en considérant la loi du maximum, **Fisher et Tippet** introduisent une transformation **affine** de  $X_{(n)}$  pour obtenir une distribution limite de  $X_{(n)}$  **non-dégénérée**<sup>6</sup> :

$$G_{(n)} := \frac{X_{(n)} - \mu_n}{\sigma_n}$$

où

---

2. Rares

3. Extreme Value Theory.

4. On considère ici  $n$  variables aléatoires **indépendantes et identiquement distribuées**  $X_1, \dots, X_n$  de variable parent  $X$  de fonction de distribution  $F_X$ .

5. Centrée et réduite.

6. On parle de loi **max-stable**.

- $(\mu_n)$  est une suite de paramètre de localisation
- $(\sigma_n) > 0$  est une suite de paramètres d'échelle

Dans le cas du **théorème central limite**, nous pouvons faire le parallèle en prenant

$$\begin{cases} \mu_n = \mathbb{E}[X_i] \\ \sigma_n = n^{-\frac{1}{2}} \sqrt{\text{Var}(X_i)} \end{cases}$$

Gnedenko montre que sous l'hypothèse d'**existence** des paramètres de normalisation, la variable du **maximum standardisée**  $G_{(n)}$  converge en loi vers une distribution **non-dégénérée**.

Le théorème suivant de Fisher-Tippett caractérise cette loi **limite**.

**Théorème 3.2.1 (Théorème de Fisher-Tippett)**

Soient  $n$  variables aléatoires **indépendantes et identiquement distribuées**  $X_1, \dots, X_n$ .

Supposons qu'il existe deux suites<sup>7</sup>  $(a_n)$  et  $(b_n)$  telles que

$$\forall x \in \mathbb{R} \quad \lim_{n \rightarrow +\infty} \mathbb{P}\{G_{(n)} \leq x\} = \lim_{n \rightarrow +\infty} \mathbb{P}\left\{\frac{X_{(n)} - \mu_n}{\sigma_n} \leq x\right\} = G(x)$$

où la distribution  $G$  est **non-dégénérée**.

La distribution **limite**  $G$ , appelée loi des valeurs extrêmes, appartient à l'une des familles<sup>8</sup> de distribution **paramétrique** de paramètre  $\xi$ <sup>9 10</sup> suivante  $G_\xi$  :

1. Famille de Fréchet

$$\exists \xi > 0 \quad G_\xi(x) := \begin{cases} \exp\{-x^{-\frac{1}{\xi}}\} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

2. Famille de Gumbel

$$\forall x \in \mathbb{R} \quad G_0(x) := \exp\{-e^{-x}\}$$

3. Famille de Weibull

$$\exists \xi > 0 \quad G_\xi(x) := \begin{cases} \exp\{-(-x)^{\frac{1}{\xi}}\} & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{cases}$$

---

7. S'il existe des suites de paramètres de standardisation  $(\mu_n)$  et  $(\sigma_n)$  telles que la distribution limite de la variable  $X_{(n)}$  standardisée est **non-dégénérée** alors on dit que  $F_X$  appartient au **domaine d'attraction** de la loi limite. On dit aussi que la distribution **limite** est un **attracteur** de  $F_X$ .

8. Extreme value (EV) distributions.

9. Le paramètre  $\xi$  est un **paramètre de forme** de la **queue** de distribution. Plus  $\xi$  est **grand** en valeur absolue, plus la queue de la loi limite  $G_\xi$  est **épaisse**. Ainsi, le poids des événements extrêmes dans la distribution initiale  $F_X$  est plus important.

10. On appelle **indice de queue** le paramètre  $\frac{1}{\xi}$ .

Le théorème de **Fisher-Tippet** **ne** garantie **pas** l'existence des suites de paramètres standardisation  $(\mu_n)$  et  $(\sigma_n)$ . De plus, s'il y a existence, la loi limite n'est pas **unique**. Néanmoins, le paramètre  $\xi$  **ne** dépend **pas** de ces suites de paramètres standardisation mais **seulement** de la fonction de distribution  $F_X$  de la variable parent  $X$ . Cette caractéristique sera utile dans le cadre de l'inférence statistique.

En **pratique**, la difficulté est de choisir la distribution extrême la plus appropriée aux **données**.

Mais, grâce aux travaux de **Von Mises** et **Jenkinson**, les distributions extrêmes de **Fréchet**, **Gumbel** et **Weibull** peuvent être combinées en un **seul** type de distribution. C'est l'objet du théorème suivant.

### **Théorème 3.2.2 (Représentation de Jenkinson-Von Mises - Distributions GEV)**

*Les lois de **Fréchet**, de **Gumbel** et de **Weibull** sont des **cas particuliers** de la loi des **extremums généralisés**<sup>11</sup> de paramètre de **localisation**  $\mu \in \mathbb{R}$ , de paramètre d'**échelle**  $\sigma > 0$  et de paramètre de **forme**  $\xi \in \mathbb{R}$  :*

$$\forall x \in \mathbb{R} \text{ tel que } 1 + \xi \frac{x - \mu}{\sigma} > 0 \quad G_{\mu, \sigma, \xi} := \begin{cases} \exp\left\{-\left(1 + \xi \frac{x - \mu}{\sigma}\right)^{\frac{-1}{\xi}}\right\} & \text{si } \xi \neq 0 \\ \exp\left\{e^{-\frac{x - \mu}{\sigma}}\right\} & \text{si } \xi = 0 \end{cases} \quad (3.1)$$

On peut *réécrire* le théorème de **Fisher-Tippet** : si  $F_X$  appartient au domaine d'attraction d'une distribution non-dégénérée  $G$  alors  $G$  est une distribution GEV.

### **3.2.2 Méthode POT - Peaks-Over-Threshold**

Nous venons de voir que la GEV est l'**unique** loi de probabilité du maximum (normalisé) d'un échantillon. Pour **estimer** les paramètres de cette GEV on utilise souvent la méthode dite par **blocs** qui consiste à créer des échantillons de maximum :

$$\begin{aligned} X_{(n)} &= \max\{X_1, \dots, X_n\} \\ &= \max\{X_{(n)[1]}, \dots, X_{(n)[m]}\} \end{aligned}$$

où  $X_{(n)[k]} = \max_{X_i \in \text{bloc } k} \{X_1, \dots, X_n\}$ .

Néanmoins cette méthode peut conduire à une perte d'information. Certains blocs peuvent contenir plusieurs extrêmes et d'autres aucun. Ainsi, pour palier à cet inconvénient nous allons présenter la méthode dite POT basée sur la loi de **Pareto généralisée**.

---

11. Generalised Extreme Value (GEV) distribution.

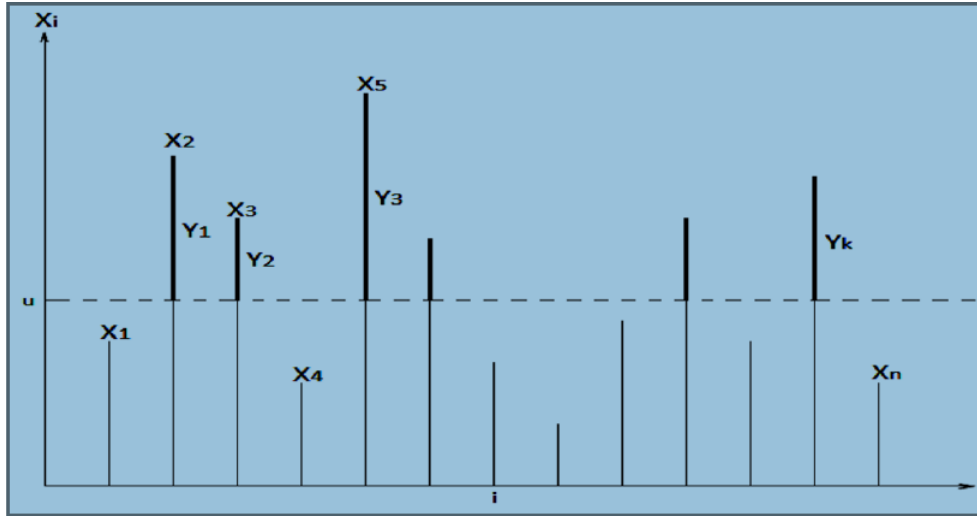


FIGURE 3.1 – L'échantillon  $X_1, \dots, X_n$  et les excès associés.

Soit  $F_u$  la loi des excès<sup>12</sup> au-delà d'un seuil  $u$  non-aléatoire. Elle est définie<sup>13</sup> par :

$$\forall 0 \leq y \leq x_F - u \quad F_u(y) := \mathbb{P}\{X - u \leq y \mid X > u\} = \frac{F(y + u) - F(u)}{1 - F(u)}$$

où  $x_F$  est le point extrême<sup>14</sup> droit défini par  $x_F := \sup\{x \in \mathbb{R} : F(x) < 1\}$ .

Ainsi on a

$$\forall 0 \leq y \leq x_F - u \quad F(y + u) = (1 - F(u))F_u(y) + F(u)$$

Si on approxime  $F(u)$  par la distribution **empirique** de la distribution  $F$

$$1 - F(u) = \mathbb{P}\{X > u\} \simeq \frac{N_u}{n} := 1 - F_n(u)$$

où  $N_u$  est le nombre d'observation dans la queue droite<sup>15</sup> i.e.  $N_u := \{i \in [1, n] \cap \mathbb{N} : x_i > u\}$

On a

$$\forall x \geq u \quad F(x) \simeq (1 - F_n(u))F_u(x - u) + F_n(u) \quad (3.2)$$

**Balkema, de Haan and Pickands** ont démontré que la distribution **asymptotique** de  $F_u$  est la distribution

12. On appelle excès de la variable  $X$  **au-delà d'un seuil**  $u < x_F$  la variable  $Y = X - u \mid X > u$  qui prend ses valeurs sur  $]0, x_F - u[$ .

13. D'après le théorème de Bayes.

14. The right end-point.

15. Upper tail.

de Pareto généralisée<sup>16 17</sup> définie par

$$\forall x \geq u \quad H_{\xi, \sigma}(x) := \begin{cases} \left(1 - \left(1 + \xi \frac{x-u}{\sigma}\right)^{-\frac{1}{\xi}}\right) 1_{\{1 + \xi \frac{x-u}{\sigma} > 0\}} & \text{si } \xi \neq 0 \\ 1 - e^{-\frac{x-u}{\sigma}} & \text{si } \xi = 0 \end{cases} \quad (3.3)$$

### **Théorème 3.2.3 (Théorème de Pickands et de Balkema-De Haan)**

Soient  $F$  la fonction de distribution d'une variable aléatoire et  $F_u$  sa distribution des excès pour un seuil  $u$  **non-aléatoire élevé**.

Pour tout  $\xi \in \mathbb{R}$ , la distribution  $F_u$  peut être **approximée**<sup>18</sup> par une distribution de Pareto généralisée  $H_{\xi, \sigma(u)}$  :

$$\lim_{u \rightarrow x_F} \sup_{0 < y < x_F - u} |F_u(y) - H_{\xi, \sigma}(y + u)| = 0$$

si et seulement si  $F$  **appartient** au domaine d'attraction d'une distribution GEV  $G_\xi$  ayant le **même** paramètre de forme  $\xi$  que celui la distribution GPD  $H_\xi$ .

Ce théorème indique que si la distribution du maximum *normalisé* converge vers une loi non-dégénérée de paramètre  $\xi$  alors la loi de **Paréto** généralisé de même paramètre  $\xi$  modélise la loi excédents au delà d'un seuil non aléatoire  $u$  suffisamment élevé.

On peut ainsi approximer la distribution  $F$  définie en (3.2) pour  $u$  élevé par :

$$\forall x \geq u \quad F(x) \simeq (1 - F_n(u))H_{\xi, \sigma}(x) + F_n(u)$$

En particulier,  $\xi \neq 0$  alors on a l'approximation

$$\boxed{\forall x \geq u \quad F(x) \simeq 1 - \frac{N_u}{n} \left(1 + \xi \frac{x-u}{\sigma}\right)^{-\frac{1}{\xi}}} \quad (3.4)$$

### **3.2.3 Estimation de la distribution de Pareto généralisée GPD**

Nous allons nous focaliser sur les lois de distributions de la famille de **Fréchet**.

Pour pouvoir estimer les paramètres  $\xi$  et  $\sigma$  de la distribution POT (Cf. le résultat 3.4) nous devons déterminer **préalablement** le seuil  $u$ .

Nous pouvons dès à présent établir quelques remarques sur le choix du seuil :

- Choisir un  $u$  trop **grand** implique trop **peu** d'observations d'excès et donc une estimation des paramètres de mauvaise qualité.
- Choisir un  $u$  trop **petit** introduit un **biais** dans l'adéquation de la GPD estimée à la queue de distribution empirique. la méthode POT étant un résultat **asymptotique**.

16. Generalized Pareto Distribution (GPD).

17. Ce théorème montre le lien qui existe entre le max-domaine d'attraction et le comportement limite de la loi des excès

18. On a une convergence uniforme.

Dans l'annexe page 167 nous présentons des méthodes graphiques pour déterminer le seuil  $u$  :

- Graphique de l'espérance excédentaire
- Graphique de stabilité Hill, Pickands, ...

Une fois le seuil **déterminé**, on peut utiliser

- un **estimateur** non-paramétrique pour déterminer le paramètre  $\xi$
- un **estimateur** paramétrique pour déterminer le paramètre de dispersion  $\sigma$

L'annexe page 167 décrit la construction de ces estimateurs.

### 3.2.4 Loi mélange

Cette étape permet d'obtenir des trajectoires cohérentes **basées** sur le **passé** dans la majorité des cas (corps de la fonction) et de néanmoins pouvoir simuler des **nouvelles** réalisations **extrêmes**.

Nous introduisons des événements **extrêmes** dans la procédure de *bootstrap* en définissant une loi mélange des résidus selon l'algorithme suivant :

**étape 1** sélectionner un modèle économétrique

**étape 2** estimer le modèle sur les données historiques

**étape 3** calculer les résidus

**étape 4** classer dans l'ordre croissant les résidus observés et choisir les seuils à gauche et à droite<sup>19</sup>.

**étape 5** estimer les distribution généralisées de Pareto pour chaque seuil.

**étape 6** effectuer un tirage aléatoire de probabilité  $n^{-1}$  d'un résidu observé

- si l'observation est supérieure au seuil à droite, générer une réalisation GPD et retourner le seuil à droite plus la variable d'excès.
- si l'observation est inférieure au seuil à gauche alors générer une réalisation GPD et retourner le seuil à gauche - variable d'excès.

---

19. Lower and upper thresholds

- si l'observation appartient au corps (coeur) de la distribution des résidus alors retourner l'observation elle-même.

**étape 7** c'est un tirage avec remise. Ajouter ce nouveau résidu à l'échantillon et répéter l'étape 6.

Ainsi, en utilisant

- l'approximation en **queue** de distribution obtenue par la méthode POT
- la distribution **empirique** pour la partie centrale, corps de la distribution

La distribution mélange<sup>20</sup> des résidus issue de cet algorithme s'écrit :

$$\widehat{F}_Z(z) = \begin{cases} \frac{k_2}{n} \left(1 + \xi_{k_2}^{(2)} \frac{z - z_{(n-k_2)}}{\beta_{k_2}^{(2)}}\right)^{-\frac{1}{\xi_{k_2}^{(2)}}} & \text{si } z < z_{(n-k_2)} \\ n^{-1} \sum_{i=1}^{n-k_2-k_1} 1_{\{z_{(n-k_2)} \leq z_i \leq z\}} & \text{si } z_{(n-k_2)} \leq z \leq z_{(k_1+1)} \\ 1 - \frac{k_1}{n} \left(1 + \xi_{k_1}^{(1)} \frac{z - z_{(k_1+1)}}{\beta_{k_1}^{(1)}}\right)^{-\frac{1}{\xi_{k_1}^{(1)}}} & \text{si } z > z_{(k_1+1)} \end{cases} \quad (3.5)$$

où les paramètres

- $k_1$  est le nombre d'observations extrêmes au-delà du seuil à droite.
- $k_2$  est le nombre d'observations extrêmes au-delà du seuil à gauche.
- $z_{(n-k_2)}$  est le seuil à gauche.
- $z_{(k_1+1)}$  est le seuil à droite.

Dans la partie centrale  $\frac{n-k_1-k_2}{n}$  représente la **proportion** des résidus centraux :

$$\frac{n - k_1 - k_2}{n} \frac{1}{n - k_1 - k_2} \sum_{i=1}^{n-k_2-k_1} 1_{\{z_{(n-k_2)} \leq z_i \leq z\}} = \frac{1}{n} \sum_{i=1}^{n-k_2-k_1} 1_{\{z_{(n-k_2)} \leq z_i \leq z\}}$$

---

20. Nous appellerons cette loi "bootstrap extrême" dans la suite du document.

## Chapitre 4

# Simulations multivariées

LORSQUE L'ON SOUHAITE SIMULER PLUSIEURS GRANDEURS ÉCONOMIQUES **SIMULTANÉMENT**, IL FAUT DOTER LE GSE D'UNE **STRUCTURE DE DÉPENDANCE** POUR PERMETTRE UNE **COHÉRENCE** DES TRAJECTOIRES SIMULÉES. C'EST L'OBJET DE CE CHAPITRE.

NOUS PRÉSENTONS

- QUELQUES STRUCTURES DE DÉPENDANCE COURAMMENT UTILISÉES DANS LES GSE.
- NOTRE STRUCTURE DE DÉPENDANCE BASÉE SUR LA COPULE DE PAUL DEHEUVELS.

NOUS OBTIENDRONS UN ALGORITHME **RÉCURSIF** GÉNÉRANT À CHAQUE PAS DES SCÉNARIOS DONT LA DISTRIBUTION EST UNE LOI **MÉLANGE MULTI-DIMENSIONNELLE** APPELÉ ICI SIMULATION *bootstrap*-EXTRÊMES.



La grande majorité des GSE utilisent une structure de dépendance **linéaire**. On rencontre, **en pratique**, trois principales classes de modèles dans les implémentations de GSE :

- les modèles **intégrés** qui utilisent des inter-dépendances linéaires issues de la théorie économique, des liens de causalité.
- les modèles **composites standards** qui utilisent la **corrélation** linéaire comme mesure de dépendance.
- les modèles **composites non-standards** qui utilisent un **structure** de dépendance plus **riche**.

## 4.1 Modèles intégrés

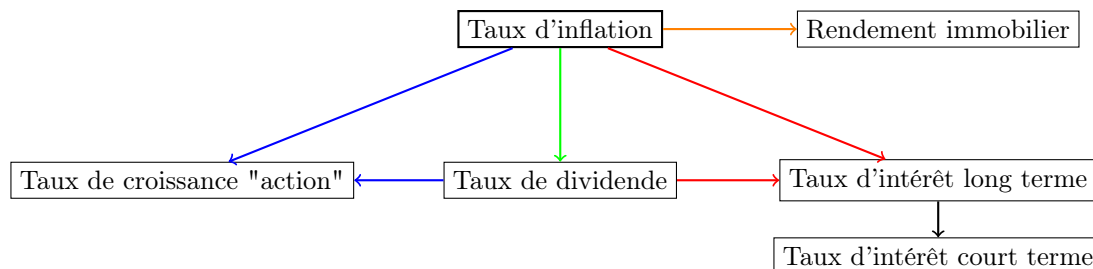
Les générateurs de scénarios économiques avec un modèle **intégré** modélisent une ou plusieurs quantités de **référence** pour **déduire** les autres grandeurs économiques. On les qualifie de GSE avec **structure en cascade**.

Dans ces modèles, la structure de dépendance est construite à partir d'une modélisation issue de **théorie économique** où l'inter-dépendance des classes d'actifs est forte. L'objectif de ces types de modèles est de pouvoir générer des trajectoires **probables** au sens économique.

Nous présentons ci-dessous la structure en cascade des principaux modèles sans décrire les dynamiques<sup>1</sup> des actifs modélisés.

### 4.1.1 Modèle de Wilkie - 1986 - 1995

Wilkie dans son article initial (Cf. [39]<sup>2</sup>) décrit un modèle de générateur de scénarios économiques avec une structure en cascade définie à partir de quatre variables dont l'**unique** variable de **référence** est l'inflation. Les dynamiques des autres variables taux de dividende des indices actions, taux d'intérêt et prix de l'immobilier sont donc décrites à partir de la dynamique du **taux d'inflation**.



Ce n'est pas une structure **multivariée** au sens strict où l'on peut décrire directement l'influence des variables l'une par rapport à l'autre par une mesure de corrélation ou par la distribution jointe. La dépendance est ici construite via la structure en cascade. Par exemple, le taux de dividende à la date  $t$  est défini à partir du

1. Ces dynamiques ont fait l'objet de description dans le paragraphe "Choix économétriques pour le GSE" page 102.  
 2. L'article de 1995 est une mise à jour de des travaux de 1986 notamment avec une nouvelle modélisation de la dynamique du taux d'inflation avec une structure ARCH.

niveau d'inflation courant et de ses valeurs passées via un processus auto-régressif. La dépendance est définie par la **sensibilité** du taux de dividende au taux d'inflation.

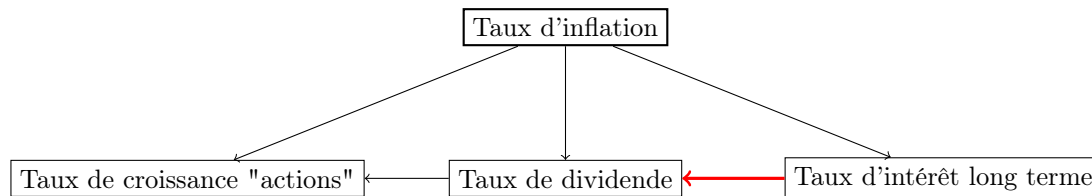
Ce modèle a été **innovateur**, il intègre pour la première fois la **construction** d'un **Générateur de Scénarios Économiques**. Les études de générations de scénarios se focalisaient **auparavant** plutôt sur une **seul** classe d'actif, **indépendamment** les unes des autres. Il introduit ainsi des **interactions**<sup>3</sup> entre plusieurs grandeurs économiques permettant une **cohérence** entre les différentes variables simulées.

Néanmoins, malgré sa relative<sup>4</sup> simplicité d'implémentation, plusieurs **critiques**<sup>5</sup> ont été apportées à ce modèle. Plusieurs études statistiques ont montré :

- une dépendance des résidus
- une non-normalité des résidus
- une hétéroscédasticité des résidus
- une non-stationnarité des résidus.

Cela suggèrent<sup>6</sup> une **non-linéarité** des dynamiques<sup>7</sup>. Des extensions ont donc été apportées pour déterminer la dynamique la plus adéquate (structure ARCH, processus TAR, ...).

**Mais dans cette section ce qui nous intéresse**, c'est la structure de dépendance **induite** par le modèle. Certaines études ont montrées par exemple que les corrélations observées entre les actions et les taux d'intérêt long terme ne sont pas reproduites par le modèle de Wilkie. Certains auteurs<sup>8</sup> ont donc proposé de modifier la structure en cascade de Wilkie et plus particulièrement en modifiant le lien de **causalité** entre les variables :



#### 4.1.2 Autres modèles en cascade

D'autres modèles ont suivi le modèle de Wilkie. Ces modèles diffèrent principalement du **choix** de la ou des variables de référence, par exemple :

---

3. Les travaux de Fama ont montré de fortes corrélations entre l'inflation et les autres actifs.  
4. Néanmoins de nombreux paramètres sont à estimer.  
5. Cf. une revue dans le mémoire de G. Rambaruth [68].  
6. Une mauvaise spécification des dynamiques : séries non-stationnaires  
7. Les dynamiques originales étaient de type auto-régressive.  
8. Cf. l'article de Whitten and Thomas.

- Le modèle de Kaufmann & al.<sup>9</sup> en 2001 présente une structure en cascade telle que les auteurs choisissent comme variable de **référence** le taux d'intérêt **court terme**.

Ce type de modèle n'a pas vraiment d'application en assurance-**vie**, les projections étant à **long** terme. En assurance **non-vie** où le risque est souvent **annuel** (sauf par exemple en responsabilité civile) ce type de modèle pourrait avoir du sens.

- Le modèle de Brennan et Xia en 2000 présente une structure en cascade avec deux quantités **deux** quantité de référence, l'**inflation** et le taux d'intérêt **réel**<sup>10\*</sup>.

Contrairement au modèle de Wilkie, l'**immobilier** n'est **pas** modélisé dans ce modèle. Or l'allocation des entreprises d'assurance est composé d'une part **non** négligeable d'actifs immobiliers.

Pour ces modèles également, la structure **dépendance** est définie par des liens de **causalité** intégrés dans les dynamiques via des liaisons (log)-**linéaires successives** induites par la théorie économique. Les simulations issues de ces modèles **intégrés** apparaissent donc **cohérentes**. Néanmoins il est souvent difficile d'extraire la structure de dépendance indépendamment des dynamiques de chaque variable considérée. Il est donc difficile d'intégrer des dynamiques plus riches. C'est ainsi qu'on été introduit les modèles composites dont la structure de dépendance s'intègre **a posteriori**, les choix des dynamiques étant effectués indépendamment.

## 4.2 Modèles composites

Dans les modèles **composites**, on **choisit** une dynamique **adaptée** à chaque grandeur économique, par exemple :

- Modèle **Black & Scholes** pour les indices actions
- Modèle **CIR** pour les taux d'intérêt
- Modèle **Vasicek** pour l'inflation
- ...

Ensuite, on dote le modèle composite d'une **structure de dépendance** qui est souvent déterminée par les **corrélations linéaires**<sup>11</sup> *historiques* entre ces différentes dynamiques.

Dans les modèles composites, les choix de dynamique des grandeurs économiques sont donc effectués de façon **indépendants**. Ils permettent donc pour **chaque** grandeur économique d'utiliser des modèles de diffusion plus ou moins complexes en fonction de l'usage qui sera fait du modèle et de l'horizon de projection.

---

9. Cf. l'article [23].

10. Taux d'intérêt net d'inflation.

11. Matrice de corrélation.

### 4.2.1 Modèle d'Ahlgrim et al. 2005

Dans leur article [3], les auteurs<sup>12</sup> proposent un modèle de GSE monde **réel** basé sur les **corrélations**. Ils conservent une structure **économique** mais sans l'intégrer directement dans les dynamiques.

Ainsi, ils choisissent trois variables de **référence** :

- Le taux d'inflation  $i_t$
- Le taux d'intérêt **réel** à long terme  $l_t$
- Le taux d'intérêt **réel** à court terme  $r_t$

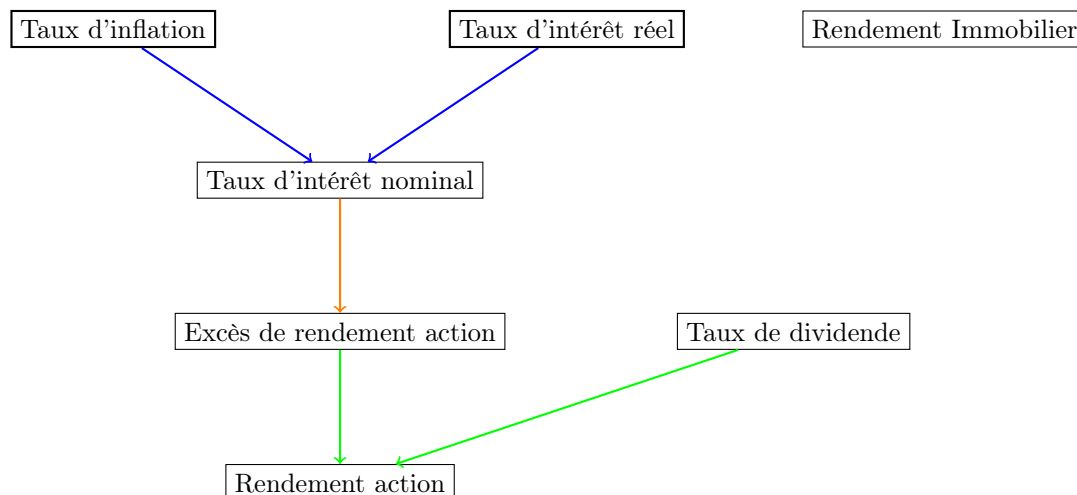
A partir de ces variables et de l'**excès** de rendement d'un investissement action, les auteurs déduisent le rendement action **hors** dividende :

$$s_t = r_t + i_t + x_t$$

Pour obtenir le rendement **total** d'un investissement action on ajoute finalement la variable **dividende** :

$$\tilde{s}_t = s_t + d_t$$

Ci-dessous, la structure économique du modèle d'Ahlgrim.



A noter que ce **n'est pas** une structure **intégrée** dans le sens où la dépendance **n'est pas incluse** dans les dynamiques<sup>13</sup>. Chaque dynamique est **indépendante**. Ensuite on utilise une structure de dépendance définie par la matrice de corrélation des résidus de chaque dynamique. Les relations simples du début du paragraphe sont utilisées pour déterminer enfin les prix des différents actifs.

12. Soutenus par les associations professionnelles américaines Casualty Actuarial Society et Society of Actuaries.

13. Pour l'inflation et le rendement immobilier : processus d'Orstein-Uhlenbeck. Pour l'excès de rendement action : modèle à changement de régime. Pour les taux d'intérêt réel : modèle de Vasicek à 2 facteurs. Cf. le paragraphe 6 page 102.

L'algorithme de simulation s'effectue en quatre étapes :

1. Détermination des lois des variables :
  - discrétisation ses différentes dynamiques.
  - estimation historique des paramètres des processus.
2. Détermination de la structure de dépendance :
  - détermination des résidus des processus
  - estimation de la matrice de corrélation des résidus.
3. Génération de loi multi-dimensionnelle :
  - simulation d'un vecteur gaussien
  - décomposition de Cholesky.

La critique de ce type de modèle est souvent le fait que structure de dépendance est basée sur les corrélations linéaires **historique**. Les simulations générées sont donc sensible à la **plage** historique d'**estimation** et cela pourrait donc conduire à des projections **non** pertinentes. Pour y remédier, certains introduisent des avis d'expert en modifiant la matrice de corrélation estimées mais cela semble difficile en dimension supérieure à 2.

D'autre part, l'utilisation d'une matrice de corrélation ne permet pas d'introduire des **dépendances de queues**. C'est l'objet du paragraphe suivant via une structure de dépendance basée sur les **copules**.

#### 4.2.2 Modèles à copules

Une structure de dépendance basée sur la copule **gaussienne** semble **inadaptée** lorsqu'on souhaite générer des scénarios économiques incorporant des trajectoires d'**événement extrêmes**. Or si la prise en compte des scénarios **extrêmes** dans les trajectoires générées est **insatisfaisante** cela conduit une **sous-estimation** du besoin global de solvabilité par exemple.

Ainsi, à la place de la matrice de corrélation, il peut être intéressant d'introduire une dépendance basée sur les **copules** pour enrichir la structure de dépendance du générateur de scénarios économiques, en particulier les dépendances de queues. On peut citer une application dans le cadre du mémoire de K. Armel<sup>14</sup>

#### 4.2.3 Modèle avec *bootstrap* vectoriel

La méthodologie du *bootstrap* vectoriel ne définit pas une structure de dépendance de façon **précise** mais elle consiste à **ré-échantillonner** par la méthode du *bootstrap* une structure de dépendance historique.

Pour ce faire, au lieu d'appliquer la méthodologie univariée décrite dans le chapitre (3) page 70 nous allons l'appliquer sur le **vecteur** des résidus.

---

14. Cf. [63].

La structure de dépendance est donc **implicite** dans le vecteur *bootstrapé* de la date  $t$ . On conserve ainsi la cohérence économique observée à la date  $t$  entre les variables.

#### 4.2.4 Simulations *bootstrap*-extrêmes

Dans le paragraphe (3.5) page 79, nous avons introduit en dimension 1 une procédure de simulation basée sur le *bootstrap* incorporant des événements **extrêmes**. Cette procédure a conduit à une distribution **mélange** des résidus. Nous allons maintenant généraliser cette procédure à **plusieurs** dimensions.

Nous présentons dans ce paragraphe la dimension 3 mais la généralisation à une dimension supérieure suit le même processus.

Considérons trois séries temporelles, réalisation des vecteurs aléatoires  $(X_j^1, \dots, X_j^n)$ ,  $j = 1, 2, 3$  :

- $(x_1^1, \dots, x_1^n)$
- $(x_2^1, \dots, x_2^n)$
- $(x_3^1, \dots, x_3^n)$

Nous **choisissons** pour chaque série un modèle **économétrique**. Nous **estimons indépendamment** les paramètres de chaque modèle. Nous obtenons ainsi trois séries de résidus *centrés* :

- $(\epsilon_1^1, \dots, \epsilon_1^n)$
- $(\epsilon_2^1, \dots, \epsilon_2^n)$
- $(\epsilon_3^1, \dots, \epsilon_3^n)$

On a donc les distributions **empiriques** de chaque série de résidus :

- $\hat{F}_{1,n}(\epsilon)$  pour la série des  $\epsilon_1$
- $\hat{F}_{2,n}(\epsilon)$  pour la série des  $\epsilon_2$
- $\hat{F}_{3,n}(\epsilon)$  pour la série des  $\epsilon_3$

En dimension 1 nous aurions effectué un tirage *bootstrap* sur chaque série. En dimension plus grande, nous considérons plutôt la série des **vecteurs** des résidus :

$$\left( \begin{array}{c} \epsilon_1^1 \\ \epsilon_2^1 \\ \epsilon_3^1 \end{array} \right), \dots, \left( \begin{array}{c} \epsilon_1^n \\ \epsilon_2^n \\ \epsilon_3^n \end{array} \right) := (\underline{\epsilon}^1, \dots, \underline{\epsilon}^n)$$

Nous effectuons alors un tirage *bootstrap* des vecteurs :

$$(\underline{\epsilon}^{1*}, \dots, \underline{\epsilon}^{n*})$$

Ensuite pour obtenir la distribution future des variables  $(x_1, x_2, x_3)$  on applique la méthodologie récursive décrite dans le paragraphe (3.1) page 71 en introduisant la  $i^{\text{ème}}$  composante du vecteur des résidus *bootstrap* dans le modèle économétrique de la  $i^{\text{ème}}$  variable.

En considérant les vecteurs de résidus, on conserve ainsi la **structure de dépendance** entre les différentes variables (résidus) à une date **donnée**.

Considérons maintenant, par exemple, le  $i^{\text{ème}}$  vecteur *bootstrap* correspondant à la  $k^{\text{ème}}$  date :

$$\underline{\epsilon}^{i*} := \begin{pmatrix} \epsilon_1^{i*} \\ \epsilon_2^{i*} \\ \epsilon_3^{i*} \end{pmatrix} := \begin{pmatrix} \epsilon_1^k \\ \epsilon_2^k \\ \epsilon_3^k \end{pmatrix}$$

Si  $X$  est une variable aléatoire de fonction de répartition  $F_X$  alors la variable  $F_X(X)$  est une variable aléatoire **uniforme** sur  $[0, 1]$ . En effet

$$\begin{aligned} \mathbb{P}\{F_X(X) < y\} &= \mathbb{P}\{X < F_X^{-1}(y)\} \\ &= F_X(F_X^{-1}(y)) \\ &= y \end{aligned}$$

On peut donc considérer le vecteur de variables suivant comme un vecteur de réalisations de variables aléatoires **uniformes** :

$$\begin{pmatrix} u_1^{i*} \\ u_2^{i*} \\ u_3^{i*} \end{pmatrix} := \begin{pmatrix} \widehat{F}_{1,n}(\epsilon_1^{i*}) \\ \widehat{F}_{2,n}(\epsilon_2^{i*}) \\ \widehat{F}_{3,n}(\epsilon_3^{i*}) \end{pmatrix}$$

$\widehat{F}_{j,n}(\cdot)$  étant la fonction de répartition **empirique** des résidus de la variable  $j$ .

D'autre part, si une variable  $u_j^{i*}$  est une réalisation d'une variable uniforme sur  $[0, 1]$  alors la variable  $u_j^{i*}$  peut-être vue comme une  $i^{\text{ème}}$  réalisation d'une variable aléatoire  $F_{j,Z}(Z_j^*)$ ,  $F_{j,Z}(z)$  étant la fonction de répartition de la loi mélange décrite dans le paragraphe (3.5) page 79 pour la variable  $j$  i.e. <sup>15</sup> :

$$u_j^{i*} = \widehat{F}_{j,Z}(z_j^{i*})$$

De façon équivalente nous avons

$$z_j^{i*} = \widehat{F}_{j,Z}^{-1}(u_j^{i*})$$

Ainsi

$$\begin{cases} z_1^{i*} := \widehat{F}_{1,Z_1}^{-1}(\widehat{F}_{1,n}(\epsilon_1^{i*})) \\ z_2^{i*} := \widehat{F}_{2,Z_2}^{-1}(\widehat{F}_{2,n}(\epsilon_2^{i*})) \\ z_3^{i*} := \widehat{F}_{3,Z_3}^{-1}(\widehat{F}_{3,n}(\epsilon_3^{i*})) \end{cases}$$

---

15. Rappelons la notation : l'indice  $j$  représente la  $j^{\text{ème}}$  variable d'intérêt et l'indice  $i$  la  $i^{\text{ème}}$  réalisation.

Nous pouvons donc **simuler** des variables  $Z_1, Z_2, Z_3$  à **partir** de tirages *bootstrap* des résidus en utilisant la relation :

$$Z_j := \widehat{F}_{j, Z_j}^{-1}(\widehat{F}_{j, n}(\Sigma_j^*))$$

et de façon équivalente

$$Z_j^* := \widehat{F}_{j, Z_j}^{-1}\left(\frac{R_j^*}{n}\right)$$

où  $R_j^*$  est la statistique de **rang** de la variable  $\Sigma_j^*$ .

Le GSE **bootstrap-extrême** consiste ainsi à simuler de façon **réursive** plusieurs variables d'intérêt sur un horizon défini en  $K$  **étapes** (pas de temps).

La  $k^{\text{ème}}$  **simulation** s'effectue en **deux** étapes :

**Etape 1 :** TIRAGE **BOOTSTRAP** DE  $n$  VECTEURS DE TAILLE  $d : \underline{\epsilon}^{1*}, \dots, \underline{\epsilon}^{n*}$ .

CALCUL DES **RANGS**  $r_j^{i*}$  DE CHAQUE COORDONNÉE  $\epsilon_j^{i*}$  DES VECTEURS  $\underline{\epsilon}^{i*}$

**Etape 2 :**

1. SI LA RÉALISATION  $\frac{r_j^{i*}}{n} \in ]\frac{k_j^1}{n}, \frac{k_j^2}{n}[$ , **COEUR** DE LA FONCTION DE RÉPARTITION DE LA LOI MÉLANGE, NOUS **CONSERVONS** LA RÉALISATION  $\epsilon_j^{i*}$  DE RANG  $r_j^{i*}$ . EN EFFET, NOUS AVONS DANS CE CAS

$$z_j^i := \widehat{F}_{j, Z_j}^{-1}(\widehat{F}_{j, n}(\epsilon_j^{i*})) = \epsilon_j^{i*}$$

2. SI LA RÉALISATION  $\frac{r_j^{i*}}{n} < \frac{k_j^2}{n}$  ALORS LA RÉALISATION APPARTIENT À LA QUEUE **GAUCHE**. EN PARTANT DE LA DÉFINITION DE LA COPULE EMPIRIQUE DE PAUL DEHEUVELS (1.14) PAGE 58 NOUS EFFECTUONS UN TIRAGE **ALÉATOIRE UNIFORME**  $u_j^{i*}$  TEL QUE

$$\frac{r_j^{i-1*}}{n} < u_j^{i*} \leq \frac{r_j^{i*}}{n}$$

NOUS **DÉDUISONS** LA  $i^{\text{ÈME}}$  RÉALISATION DE LA VARIABLE  $Z_j$  :

$$\frac{k_j^2}{n} \left(1 + \xi_{k_j^2}^{(2)} \frac{z_j^i - \epsilon_j^{(n-k_j^2)} - \frac{1}{\xi_{k_j^2}^{(2)}}}{\beta_{k_j^2}^{(2)}}\right) = u_j^{i*}$$

$$z_j^i = \left( \left( \frac{nu_j^{i*}}{k_j^2} \right)^{-\xi_{k_j^2}^{(2)}} - 1 \right) \times \frac{\beta_{k_j^2}^{(2)}}{\xi_{k_j^2}^{(2)}} + \epsilon_j^{(n-k_j^2)}$$

3. SI LA RÉALISATION  $\frac{r_j^{i*}}{n} > \frac{k_j^1}{n}$  ALORS LA RÉALISATION APPARTIENT À LA QUEUE **DROITE**. DE LA MÊME FAÇON, NOUS EFFECTUONS UN TIRAGE **UNIFORME** SUR LE SUPPORT  $\left[ \frac{r_j^{i*}}{n}, \frac{r_j^{i+1*}}{n} \right]$ .



NOUS DÉDUISONS LA  $i^{\text{ÈME}}$  RÉALISATION DE LA VARIABLE  $Z_j$  :

$$1 - \frac{k_1}{n} \left(1 + \xi_{k_1}^{(1)} \frac{z - \epsilon_j^{(k_1+1)}}{\beta_{k_1}^{(1)}}\right)^{-\frac{1}{\xi_{k_1}^{(1)}}} = u_j^{i^*}$$

$$z_j^i = \left( \left( \frac{n - nu_j^{i^*}}{k_j^1} \right)^{\xi_{k_1}^{(1)}} - 1 \right) \times \frac{\beta_{k_j^1}^{(1)}}{-\xi_{k_j^1}^{(1)}} + \epsilon_j^{(k_j^1+1)}$$

**Etape 3 :** POUR LA  $(k+1)^{\text{ÈME}}$  SIMULATION RÉITÉRER LA PROCÉDURE À L'ÉTAPE 1 DE FAÇON INDÉPENDANTE.

Les simulations "**bootstrap-extrême**" en 3D sont indépendantes tout en respectant la structure de dépendance induite par de la loi jointe :

$$\begin{cases} z_j^i = \epsilon_j^{i^*} & \text{si } r_j^{i^*} \in ]k_j^1, k_j^2[ \\ z_j^i = \left( \left( \frac{u_j^{i^*}}{k_j^2} \right)^{-\xi_{k_j^2}^{(2)}} - 1 \right) \times \frac{\beta_{k_j^2}^{(2)}}{\xi_{k_j^2}^{(2)}} + \epsilon_j^{(n-k_j^2)} & \text{avec } u_j^{i^*} \in \left] \frac{r_j^{i^*-1}}{n}, \frac{r_j^{i^*}}{n} \right] \quad \text{si } r_j^{i^*} < k_j^2 \\ z_j^i = \left( \left( \frac{n - nu_j^{i^*}}{k_j^1} \right)^{\xi_{k_j^1}^{(1)}} - 1 \right) \times \frac{\beta_{k_j^1}^{(1)}}{-\xi_{k_j^1}^{(1)}} + z_j^{(k_j^1+1)} & \text{avec } u_j^{i^*} \in \left[ \frac{r_j^{i^*}}{n}, \frac{r_j^{i^*+1}}{n} \right[ \quad \text{si } r_j^{i^*} > k_j^1 \end{cases}$$

La loi jointe (3.5) page 79 peut donc être approximée empiriquement par la copule de Paul Deheuvels, en effet<sup>16</sup> :

$$\begin{aligned} \mathbb{P}\{Z_1 \leq z_1, Z_2 \leq z_2, Z_3 \leq z_3\} &= \mathbb{P}\left\{ \Sigma_1^* \leq \widehat{F}_{1,n}^{-1}(\widehat{F}_{1,Z_1}(z_1)), \Sigma_2^* \leq \widehat{F}_{2,n}^{-1}(\widehat{F}_{2,Z_2}(z_2)), \Sigma_3^* \leq \widehat{F}_{3,n}^{-1}(\widehat{F}_{3,Z_3}(z_3)) \right\} \\ &\simeq n^{-1} \sum_{i=1}^n \prod_{j=1}^3 1_{\epsilon_j^{i^*} \leq F_{j,n}^{-1}(\widehat{F}_{j,Z_j}(z_j))} \\ &\simeq n^{-1} \sum_{i=1}^n \prod_{j=1}^3 1_{F_{j,n}(\epsilon_j^{i^*}) \leq \widehat{F}_{j,Z_j}(z_j)} \\ &\simeq n^{-1} \sum_{i=1}^n \prod_{j=1}^3 1_{\frac{r_j^{i^*}}{n} \leq \widehat{F}_{j,Z_j}(z_j)} \end{aligned}$$

Il s'agit de la copule de Paul Deheuvels :

$$C\left(\frac{t_1}{n}, \frac{t_2}{n}, \frac{t_3}{n}\right) = n^{-1} \sum_{i=1}^n \prod_{j=1}^3 1_{r_j^{i^*} \leq t_j}$$

16. Les  $\Sigma_j^*$  sont des variables aléatoires dont les réalisations sont les  $\epsilon_j^{i^*}$ .

aux points  $t_j = n \widehat{F}_{j, Z_j}(z_j)$ .

Ainsi, en travaillant sur les **rangs** dans la procédure de simulation page 88 on **conserve** les liaisons dans les extrêmes des différents résidus **induite** par la copule empirique.

## Chapitre 5

# Modélisation économétrique

LA **PREMIÈRE** ÉTAPE DE NOTRE ALGORITHME DE SIMULATION EST LA DÉTERMINATION DU MODÈLE ÉCONOMÉTRIQUE **ADÉQUAT** ET SON ESTIMATION POUR POUVOIR CONSTRUIRE L'ÉCHANTILLON DE *Bootstrap* CONSTITUÉ DES RÉSIDUS.

L'OBJET DE CE CHAPITRE EST DE DÉCRIRE QUELQUES <sup>1</sup>

- MODÈLES ÉCONOMÉTRIQUES
- TESTS STATIQUES POUR VALIDER LA BONNE QUALITÉ DES MODÈLE ÉCONOMÉTRIQUE
- MÉTHODES D'ESTIMATION POUR DÉTERMINER LES PARAMÈTRES

---

1. Pour une description plus complète, nous renvoyons le lecteur aux ouvrages de référence [40] et [41] pour les modèles linéaires et le cours de Charpentier [43] pour une présentation en français des modèles linéaires et non-linéaires.

## 5.1 Modèles de séries temporelles

Ce chapitre présentera succinctement\* plusieurs modèles économétriques **candidats** à la modélisation de séries temporelles financières et économiques.

### 5.1.1 Bruit blanc

Un processus  $X_t$  est **stationnaire de second ordre**<sup>2</sup> si les trois conditions suivantes sont satisfaites :

1.  $\mathbb{E}[X_t^2] < +\infty$
2.  $\forall t \in \mathbb{Z} \quad \mathbb{E}[X_t] = \mu$  **indépendance** par rapport à  $t$
3.  $\forall t \in \mathbb{Z} \quad \forall h \in \mathbb{Z} \quad \text{Cov}(X_t, X_{t-h}) =: \gamma_h$  **indépendance** par rapport à  $t$

Un exemple de processus stationnaire de second ordre est le **bruit blanc**.

Un **bruit blanc faible**  $\{\epsilon_t, t \in \mathbb{Z}\}$  est un processus **stationnaire** de second ordre, tel que les auto-covariances sont toutes **nulles** :

$$\begin{aligned}\gamma_h &= 0 \quad \forall h \neq 0 \\ \gamma_0 &= \sigma^2 \quad (h = 0)\end{aligned}$$

Remarquons qu'on ne parle pas d'hypothèse d'**indépendance**. L'hypothèse d'indépendance est rarement vérifiée en particulier pour les séries financières<sup>3</sup>.

Lorsque l'hypothèse d'indépendance est vérifiée on parle alors de **bruit blanc fort**.

Un **bruit blanc gaussien** est un bruit blanc fort de loi **gaussienne**. Cela correspond à une stationnarité **stricte** i.e. stabilité de la loi. Pour les séries financières, on considère plutôt la stationnarité de second ordre plus restrictive, en effet on observe souvent des lois à queues épaisses donc avec peu de moments.

Le bruit blanc est un processus important dans la théorie des séries temporelles. Il permet de construire des processus plus **complexe** comme nous le verrons dans les paragraphes suivants.

Rappelons auparavant le théorème de représentation Wold où tout processus **stationnaire** de second ordre peut s'écrire<sup>4</sup> comme une **somme**<sup>5</sup> pondérée de chocs aléatoires orthogonaux, centrés et de même variance :

$$X_t = \epsilon_t + \sum_{i=1}^{+\infty} \phi_i \epsilon_{t-i}$$

---

2. Ou stationnaire **faible**.

3. On peut observer par exemple des auto-corrélations des résidus nulles et les auto-corrélations des carrés des résidus non-nulles. Cf. les processus ARCH.

4. En tronquant cette équation on montre que l'ensemble des moyennes mobiles finies est dense dans l'ensemble des processus stationnaires de second ordre.

5. Éventuellement augmentée d'une composante déterministe.

ainsi on peut écrire en prenant l'espérance conditionnelle

$$X_t = \epsilon_t + \mathbb{E}\left[X_t \mid \epsilon_{t-1}, \epsilon_{t-2}, \dots\right]$$

Cette équation s'interprète que tout processus stationnaire s'écrit comme la somme d'une composante **connue** à la date  $t-1$  **et** d'un choc (innovation linéaire) **nouveau** à la date  $t$ .

Dans les séries **financières** et **économiques**, nous cherchons **donc** à obtenir des processus **stationnaires**.

### 5.1.2 Processus auto-régressifs - (AR)

Un processus **auto-régressif** d'ordre 1<sup>6</sup> noté AR(1) est un processus stationnaire de second ordre de la forme :

$$X_t = \phi X_{t-1} + \epsilon_t \quad \forall t \in \mathbb{Z} \quad (5.1)$$

tel que

- $\phi$  est une constante ( $\neq 0$ )
- $(\epsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc de variance  $\sigma^2$ .
- La moyenne  $\mu$  du processus est nulle sinon remplacer  $X_t$  par  $(X_t - \mu)$  i.e.

$$X_t - \mu = \phi(X_{t-1} - \mu) + \epsilon_t \Leftrightarrow X_t = \alpha + \phi X_{t-1} + \epsilon_t \text{ avec } \alpha := 1 - \phi$$

Si le coefficient  $\phi \neq \pm 1$ , il existe un **unique** processus **stationnaire** tel que

$$(1 - \phi L) X_t = \epsilon_t$$

et on peut **inverser** le polynôme  $(1 - \phi L)^{-1}$  pour obtenir le processus  $X_t$  en fonction des  $\{\epsilon_t\}$ .

C'est ainsi que la recherche de la **racine unitaire** est à la base des tests de **stationnarité**<sup>7</sup>.

On définit la fonction d'**auto-corrélation** d'un processus stationnaire par :

$$\rho(h) = \frac{\gamma_h}{\gamma_0} =: \text{Corr}(X_{t-h}, X_t)$$

Pour un processus AR(1) on a la relation de **Yule-Walker** :

$$\rho(h) - \phi \rho(h-1) = 0 \quad \forall h > 0$$

Ainsi, la fonction d'auto-corrélation d'un processus AR(1) s'écrit  $\rho(h) = \phi^h$ .

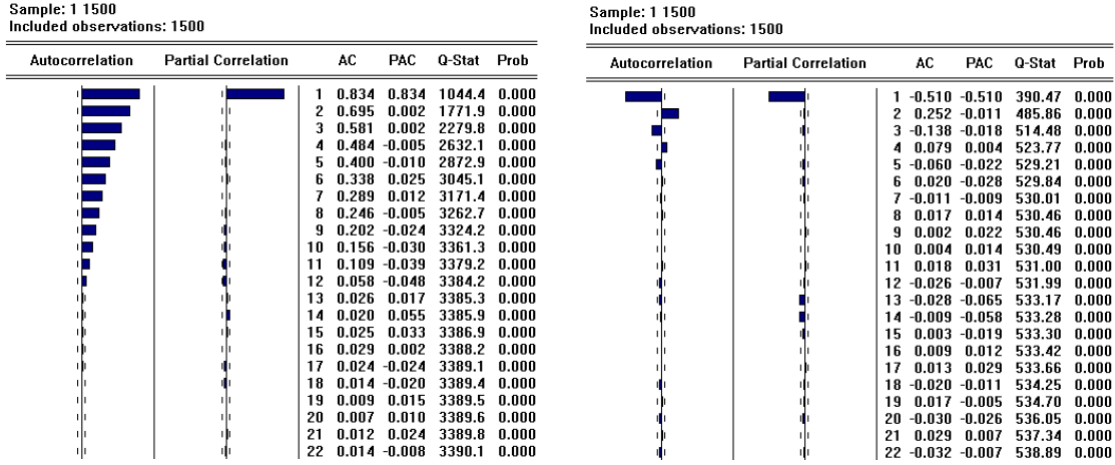
On montre ainsi que les fonctions d'auto-corrélation des processus AR(1) :

---

6. On généralise aux processus ARIMA( $p, d, q$ ) avec  $p$  l'ordre auto-régressif,  $q$  l'ordre de moyenne mobile et  $d$  l'ordre de différenciation mais nous nous attardons ici aux processus fréquemment utilisés AR(1).

7. Cf. le paragraphe page 100.

- **décroissent** exponentiellement vers  $\mathbf{0}$  avec  $0 < \phi < 1$
- **décroissent** avec une sinusoïde amortie vers  $-1 < \phi < 0$



Rappelons que l'intervalle de confiance pour un niveau de confiance des fonctions d'auto-corrélations d'un bruit blanc est <sup>8</sup> :

$$I_C = \left[ -\frac{q_{1-\alpha}}{\sqrt{H}}, \frac{q_{1-\alpha}}{\sqrt{H}} \right]$$

Dans le cadre des processus AR d'ordre  $p > 1$ , on définit également les fonctions d'auto-corrélations partielles pour le processus  $\{X_t, t \in \mathbb{Z}\}$  par :

$$r(h) := \rho(\hat{X}_t, \hat{X}_{t-h})$$

avec la projection <sup>9</sup>

$$\begin{aligned} \hat{X}_t &= X_t - \text{Proj} [X_t | X_{t-h+1}, \dots, X_{t-1}] \\ \hat{X}_{t-h} &= X_{t-h} - \text{Proj} [X_{t-h} | X_{t-h+1}, \dots, X_{t-1}] \end{aligned}$$

Ici, on regarde donc les auto-corrélations entre les variables  $X_t$  et  $X_{t-h}$  **en enlevant** les interactions avec les variables **intermédiaires**  $X_{t-h+1}, \dots, X_{t-1}$ .

On montre que les fonctions d'auto-corrélations partielles des processus AR( $p$ ) s'annulent pour un ordre supérieur  $p$  au lag  $h$  :

$$r(h) = 0 \quad \forall h > p$$

Ainsi si le **auto-corrélogramme partiel** d'une série temporelle s'**annulent** à partir d'un rang  $p + 1$  on peut envisager un modèle de type AR( $p$ ). Par exemple pour un AR(2) :

8. D'après le TCL et le théorème de Slutsky.

9. Proj désigne la **projection** dans l'espace  $L^2$  muni du produit scalaire  $\langle X | Y \rangle := \mathbb{E}[XY]$ .

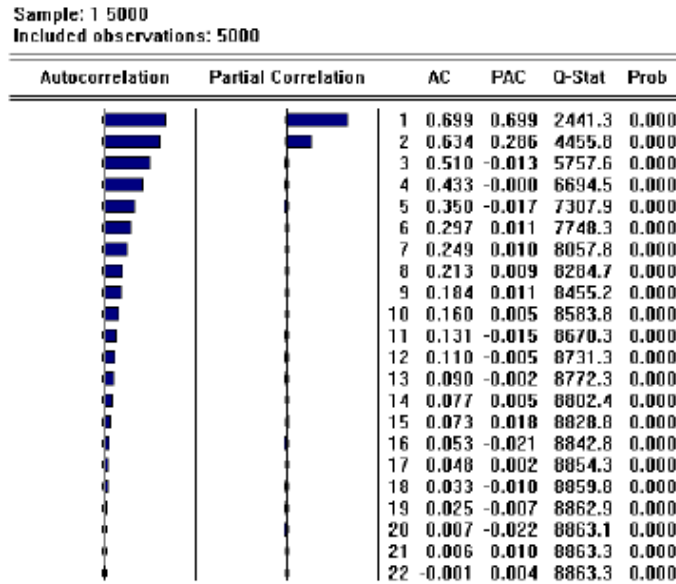


FIGURE 5.1 – Corrélogramme partiel d'un processus AR(2) avec paramètres positifs

### 5.1.3 Processus auto-régressifs conditionnellement hétéroscédastiques - (ARCH)

Nous venons de présenter les modèles linéaires de type AR. Or on observe souvent des phénomènes **non**-linéaires dans les processus financiers et économiques qui rendent les séries considérées **non**-stationnaires, y compris en ayant éventuellement supprimer la tendance ou en ayant en différencier les séries. Une modélisation **non**-linéaire devient indispensable.

Sur le graphique ci-dessus, on remarque des sous-périodes de **forte** variations des rendements et des sous-périodes de **faible** variation des rendements qui se succèdent de façon non-périodique. Ceci n'est pas incompatible avec l'homoscédasticité de la variance non-conditionnelle du processus entier. Néanmoins, on observe qu'une forte valeur du carré des rendements en  $t - 1$  apparaît augmenter la probabilité de survenance d'une forte valeur du **carré** du rendement en  $t$  (quel que soit le signe du rendement). C'est le phénomène de **regroupement**<sup>10</sup> des **extrêmes**. Ainsi la variance conditionnelle par rapport aux valeurs passées n'est pas constante c'est qu'on appelle **hétéroscédasticité conditionnelle**<sup>11</sup>.

Pour modéliser le phénomène spécifique de non-linéarité dû au caractère **hétéroscédastique** des séries, Engle dans son article [18] en 1982 et ensuite Bollerslev [9] en 1986 ont proposé une **nouvelle** catégorie de modèle permettant de modéliser les séries présentant le caractère hétéroscédastique : les processus ARCH et leur généralisation GARCH.

10. Volatility clustering.

11. C'est-à-dire qu'elle est **non**-constante dans le temps

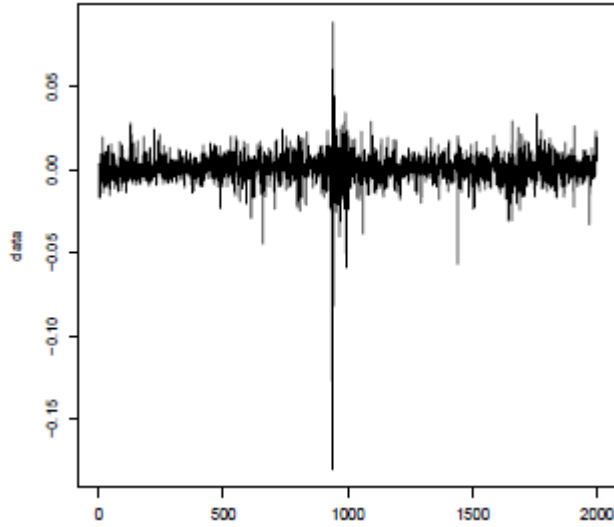


FIGURE 5.2 – Exemple d'hétéroscédasticité d'un indice boursier

Une série  $X_t$  satisfait une représentation de type ARCH(1) d'ordre 1 si

$$X_t = u_t h_t$$

avec  $h_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$  et où  $u_t$  est un bruit blanc *faible*.

La composante  $h_t^2$  désigne une variable qui **conditionnellement** à l'information constituée des valeurs passées de  $X_t$  est **déterministe** et **positive**.

L'ensemble  $\underline{X}_{t-1} := \{X_{t-1}, \dots, X_{t-H}, \dots\}$  désigne l'information **passée** et  $m_{t-1}[\cdot]$  le **moment conditionnel** par rapport à cette information

On a les moments conditionnels :

$$\mathbb{E}_{t-1}[u_t h_t] = h_t \mathbb{E}_{t-1}[u_t] = h_t \mathbb{E}[u_t] = 0$$

$$\text{Var}_{t-1}(u_t h_t) = h_t^2 \text{Var}_{t-1}(u_t) = h_t^2 \text{Var}(u_t) = h_t^2$$

Dans les processus GARCH la variance conditionnelle dépend également des valeurs **passées** de la variance  $h_t^2$ . Pour un GARCH(1, 1) nous avons :

$$h_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 h_{t-1}^2$$

Si  $X_t$  est un processus GARCH( $p, q$ ) alors<sup>12</sup>  $X_t^2$  est un processus ARIMA( $p, 0, q$ ). Cette propriété est donc utile dans l'estimation et l'identification des différents ordres.

<sup>12</sup>. Sous l'hypothèse de stationnarité des  $X_t^2$ .



### 5.1.4 Modèles avec erreurs ARCH - (ARIMA – GARCH)

Un processus est dit **auto-régressif** avec des résidus admettant une représentation ARCH si :

$$X_t = \mathbb{E}_{t-1}[X_t] + \epsilon_t \quad (5.2)$$

avec  $\epsilon_t$  un processus de type ARCH.

En particulier, un processus AR(1) – ARCH(1) est un modèle auto-régressif d'ordre 1 avec une erreur de type ARCH d'ordre 1 :

$$X_t = \phi X_{t-1} + \epsilon_t$$

avec  $\epsilon_t = u_t \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2} = u_t h_t^2$  et  $u_t$  est un bruit blanc **faible** de variance **unitaire**.

Si  $u_t$  est un bruit blanc (fort) **gaussien**, la loi conditionnelle est  $\mathcal{N}(\phi X_{t-1}, h_t^2)$

Pour un AR(1) – GARCH(1, 1) nous avons la représentation suivante :

$$X_t = \phi X_{t-1} + u_t h_t \quad \text{avec} \quad h_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + h_{t-1}^2$$

Ces processus sont adaptés aux séries dont le carré des rendements sont auto-corrélés donc particulièrement pour les séries des indices "*action*"<sup>13</sup>.

### 5.1.5 Modèles à régimes

#### 5.1.5.1 Modèles auto-régressifs à seuil - (TAR)

Les modèles **auto-régressifs à seuil** permettent de prendre en compte des phénomènes **non**-linéaires. Ils permettent de modéliser plusieurs dynamiques (**régimes**) pour une même série. Dans les modèles TAR<sup>14</sup>, le mécanisme de changement de régime est spécifiée par une variable endogène<sup>15</sup>.

Les modèles SETAR<sup>16</sup> en sont une variante. Ils permettent de modéliser les phénomènes d'asymétrie c'est-à-dire suite à la suite d'un même choc la dynamique est différente en fonction de la position de la variable passée.

Un modèle SETAR d'ordre 2 de paramètre de **rupture**  $c$  à **deux** régimes a pour écriture :

$$X_t = \begin{cases} \alpha_{11} + \alpha_{21} X_{t-1} + \epsilon_t & \text{si } X_{t-1} \leq c \\ \alpha_{21} + \alpha_{22} X_{t-1} + \epsilon_t & \text{si } X_{t-1} > c \end{cases}$$

Ce modèle peut se ré-écrire

$$X_t = (\alpha_{11} + \alpha_{21} X_{t-1}) 1_{\{X_{t-1} \leq c\}} + (\alpha_{21} + \alpha_{22} X_{t-1}) 1_{\{X_{t-1} > c\}} + \epsilon_t$$

où  $(\epsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc.

**faire une remarque sur les processus Smooth Transition Autoregression**

13. Cf. le chapitre "Choix économétrique" page 118.

14. Threshold AutoRegressive model.

15. Et d'un seuil.

16. Self-Exiting Threshold AutoRegressive model. Le terme *exciting* signifie que la variable seuil est la variable elle-même avec un décalage (*lag*).

### 5.1.5.2 Modèles auto-régressifs exponentiels - (ExpAR)

Les modèles auto-régressifs exponentiels permettent de prendre en compte les phénomènes de *cluster* de volatilité.

Un processus économétrique Soft-ExpAR<sup>17</sup> est défini pour un processus **stationnaire**  $\{X_t\}_{t \geq 0}$  par :

$$X_t = \mu + \left( \rho_1 + \rho_2 e^{-X_{t-1}^2} \right) X_{t-1} + \epsilon_t = \Phi_{\mu, \rho_1, \rho_2}^{\text{Soft-ExpAR}}(X_{t-1}) + \epsilon_t \quad (5.3)$$

où  $(\epsilon_t)_{t \in \mathbb{Z}}$  est un **bruit blanc** de variance  $\sigma_\epsilon$ .

Si l'on remplace ce bruit blanc par un processus de type ARCH on définit un processus à régime avec des résidus admettant une structure ARCH :

Un modèle économétrique de type ExpAR(1) – ARCH(1)<sup>18</sup> est défini pour un processus **stationnaire**  $\{X_t\}_{t \geq 0}$  par :

$$X_t = \mu + \left( \rho_1 + \rho_2 e^{-X_{t-1}^2} \right) X_{t-1} + \epsilon_t = \Phi_{\mu, \rho_1, \rho_2}^{\text{ExpAR-ARCH}}(X_{t-1}) + \epsilon_t \quad (5.4)$$

avec  $\epsilon_t = u_t \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2}$  et  $u_t$  est un bruit blanc **faible** de variance unitaire.

#### Hypothèse 5.1.1

On supposera que  $|\alpha_1| < 1$

Si cette hypothèse est vérifiée alors

- pour des **grandes** valeurs de  $|X_{t-1}|$  le processus défini en (5.4) se comporte **comme** un processus AR(1) avec pour coefficient autoregressif  $\alpha_1$ .
- inversement pour des **petites** valeurs ( $\simeq 0$ ) de  $|X_{t-1}|$  le processus défini en (5.4) se comporte **comme** un processus AR(1) avec pour coefficient auto-régressif  $\alpha_1 + \alpha_2$ .

## 5.2 Tests statistiques

Afin de **déterminer** et de **valider** la spécification des modèles économétriques choisis, on est amené à effectuer une batterie de **tests statistiques**. C'est l'objet de cette section<sup>19</sup>. La première partie sera consacrée à **valider** les **hypothèses** des modèles **envisageables** :

- tests de significativité
- tests d'adéquation
- tests d'auto-corrélation

---

17. Soft EXponential AutoRegressive model.

18. EXponential Threshold AutoRegressive model

19. Pour une description plus détaillée Cf. (B) page 149

- tests d’homoscédasticité
- tests de stationnarité

La seconde partie permettra une fois le sous-ensemble de modèles **admissibles** déterminés de **choisir** celui qui est le plus *adéquat* aux données observées selon un **critère** statistique :

- critère de Akaike
- critère bayésien

### 5.2.1 Tests de significativité

Considérons un échantillon historique de longueur  $H$  définie par la série  $(x_{t-H}, \dots, x_{t-1})$ .

On souhaite modéliser la dynamique de cet historique par un modèle **auto-régressif** à constante avec  $p$  retards de la variable endogène :

$$X_t = \alpha_0 + f(X_{t-1}, \dots, X_{t-p}; \alpha_1, \dots, \alpha_p) + \epsilon_t \quad \text{avec } f \text{ une fonction à variables séparables}^{20} \quad (5.5)$$

$\epsilon_t$  est un bruit **blanc**.

Les tests de **significativité** consiste à valider la **présence** de la variable  $x_{t-p}$  dans le modèle en tant que variable **explicative**.

### 5.2.2 Tests d’adéquation

On s’intéresse à la **loi** des **résidus** du modèle économétrique choisi.

Un test d’**adéquation**<sup>21</sup> permet de valider l’**hypothèse** de **distribution** des résidus.

Parmi les tests d’adéquation, la conformité à la loi normale est souvent testée du fait que l’hypothèse de normalité **sous-tend** souvent les résultats asymptotiques des tests statistiques. c’est l’objet des tests de **Shapiro-Wilk** et de **Jarque-Bera**. Mais commençons par un test général d’adéquation celui de Kolmogorov-Smirnov.

### 5.2.3 Tests d’autocorrélation sérielles des résidus

On observe souvent dans les problématiques économétriques un phénomène d’**auto-corrélation** c’est-à-dire pour lesquels les résidus **successifs** sont **auto-corrélés**, cela peut être dû à une **mauvaise** spécification du modèle de régression. Ainsi a été développé les tests d’**auto-corrélation sérielles**.

### 5.2.4 Test d’hétéroscédasticité conditionnelle

Dans ce type de test, on souhaite *vérifier* la présence d’une structure ARCH dans la chronique observée.

20. Pour prendre en considération les fonctions indicatrices par exemple.

21. Goodness-of fit tests.

### 5.2.5 Tests de stationnarité

De manière générale, les méthodologies de séries temporelles, et en particulier celles développées dans le cadre des travaux de Box et Jenkins sous-tendaient l'hypothèse **fondamentale** de **stationnarité**<sup>22</sup>.

Si la série n'est pas stationnaire, il faut donc essayer de **stationnariser** la série ensuite on peut appliquer les techniques usuelles d'estimations des séries temporelles. Si cette vérification n'est pas effectuée on peut se retrouver dans le cas, par exemple, de la problématique de régression **fallacieuse**<sup>23</sup> où la régression linéaire avec des variables non-stationnaires est **non**-valide.

Il existe plusieurs types de **non**-stationnarité :

- stationnarité **en tendance**. Par exemple,  $X_t = \mu + \lambda t + \epsilon_t$ .

On peut rendre la série est stationnaire en **supprimant** la tendance **temporelle**.

- stationnarité **en différence**. On peut donner l'exemple de la **marche aléatoire**<sup>24</sup> sans dérive  $X_t = X_{t-1} + \epsilon_t$ .

On peut rendre la série stationnaire en **différenciant**<sup>25</sup> la série.

- stationnarité **en rupture**. La non-stationnarité provient d'une rupture même s'il peut y avoir une stationnarité par morceaux. C'est l'objet du test de Chow.

Pour cette section, on peut se référer à l'article très complet de B. Salanié [35].

### 5.2.6 Critère de comparaison de modèles

L'étape d'identification du type de modèle conserve un caractère **subjectif** même si l'analyse descriptive des données et les tests statistiques peut apporter une **intuition** sur le modèle adéquat. Ainsi l'identification du *meilleur* modèle peut mener à des choix **différents** selon l'utilisateur et pour un même utilisateur mener à des modèles **concurrents**.

Par exemple, pour les modèles de type ARMA stationnaire l'objet est la recherche des ordre  $p$  et  $q$  optimaux. On peut avoir une intuition via par exemple les auto-corrélations sur les ordres maximaux, le risque est donc de choisir des ordres maximum trop petits.

Pour déterminer un critère de sélection, par exemple, sachant une trajectoire observée, on peut penser que l'on choisira le modèle pour lequel la trajectoire est la plus vraisemblable. Mais ce critère basé sur la seule vraisemblance ne peut être valable pour des modèles emboîtés, les processus AR sont des cas particulier des ARMA avec des contraintes sur les coefficients. On peut choisir également le modèle pour une même vraisemblance celui qui est le plus parcimonieux i.e. celui qui a le plus petit nombre de paramètres.

---

22. Cf. la définition (5.2.5) page 100.

23. Trouver une référence.

24. Graphiquement l'explosion de la variance n'est pas visible.

25. C'est ainsi que l'on modélise souvent les log-rendements des séries plutôt que la série elle-même. On parle de processus intégrés.

C'est ainsi que sont construits la majorité de critères de comparaison :

$$C_{H,\Theta} = \frac{-2l}{H} + \frac{1}{H} \text{card}(\Theta)p(H)$$

où

—  $l$  est la vraisemblance

—  $\text{card}(\Theta)p(\cdot)$  est une mesure de pénalité fonction croissante du nombre de paramètres ( $\Theta$ ) à estimer.

Dans cette écriture la vraisemblance est considérée négativement et la pénalité positivement la règle de décision est donc de choisir le modèle ayant le critère le plus faible. Remarquons que lorsque l'on augmente la complexité on augmente la vraisemblance.

Voici les deux principaux critères que l'on retrouve couramment dans la littérature ou dans les sorties des logiciels statistiques :

— AIC de Akaike :  $C_{H,\Theta} = \frac{-2l}{H} + \frac{\text{card}(\Theta)}{H} 2 \quad p(H) = 2$

— BIC de Schwarz :  $C_{H,\Theta} = \frac{-2l}{H} + \frac{\text{card}(\Theta)}{H} \ln H \quad p(H) = \ln H$

On peut remarquer que  $\lim_{H \rightarrow +\infty} \frac{p(H)}{H} = 0$

### 5.3 Estimation paramétrique

Une fois le modèle adéquat déterminés nous devons procéder à l'estimation des paramètres du modèle.

Nous utilisons généralement les outils statistiques suivants décrits en annexe (C) page 158 :

— Moindres carrés non-linéaires

— Quasi-Maximum de vraisemblance

## Chapitre 6

# Choix économétriques pour le GSE

NOUS AVONS INTRODUIT DANS LE CHAPITRE PRÉCÉDENT DIFFÉRENTS TYPES DE **MODÉLISATION ÉCONOMÉTRIQUE** AINSI QUE LES MÉTHODES D'ESTIMATIONS INHÉRENTES.

MAINTENANT, APPLIQUONS SUR LES DONNÉES **RÉELLES** POUR CHAQUE<sup>1</sup> VARIABLE D'INTÉRÊT (TAUX, ACTION, CRÉDIT, INFLATION ET IMMOBILIER) DE NOTRE GSE.

---

1. De façon indépendante.

## 6.1 Modèle de taux d'intérêt à paramètres dynamiques - Fonctionnelle de Nelson-Siegel

Pour estimer la structure par terme des taux d'intérêt, nous choisissons d'utiliser la fonctionnelle paramétrique de **Nelson-Siegel**. Sa paramétrisation est assez flexible pour permettre de décrire de multiples formes de structure par terme de taux d'intérêt observées. De plus, le modèle de **Nelson-Siegel** présente des caractéristiques souhaitables qui permettent d'approximer un modèle sans arbitrage.

La fonctionnelle de **Nelson-Siegel** décrivant la structure par terme des taux d'intérêt a la forme paramétrique suivante :

$$R_t(\tau) = \beta_{1t} + (\beta_{2t} + \beta_{3t})\left(\frac{1 - e^{-\frac{\tau}{\lambda}}}{\lambda}\right) + \beta_{3t}e^{-\frac{\tau}{\lambda}}$$

Dans le cadre des taux d'intérêt, nous n'allons pas supposer un modèle économétrique sur le taux d'intérêt  $R_t(\tau)$  mais faire l'hypothèse de dynamique économétrique sur les trois paramètres dépendant du temps  $\beta_{1t}, \beta_{2t}, \beta_{3t}$ .

Nous allons utiliser trois processus  $AR(1)$ <sup>2</sup> pour chaque paramètre  $\beta_i$

$$\beta_{it} = \alpha + \phi_i \beta_{it-1} + \epsilon_t$$

Dans le cadre des taux d'intérêt, introduire un modèle économétrique sur les paramètres plutôt que sur la variable elle-même permet de conserver la caractéristique de courbe de taux en AOA<sup>3</sup>.

### 6.1.1 Introduction

Avant de passer à l'étape d'estimation, nous rappelons rapidement quelques définitions sur les taux d'intérêt.

Le **taux zéro-coupon continu**  $R(t, T)$  prévalant entre la date  $t$  et la maturité  $T$  correspond au rendement **annuel** d'un investissement de  $P(t, T)$  unités monétaires à la date  $t$  permettant d'obtenir **une** unité monétaire à la date  $T$  tel que les intérêts s'accroissent **continûment** :

$$P(t, T)e^{R(t, T)\tau(t, T)} = 1$$

ou encore

$$P(t, T) = e^{-\tau(t, T)R(t, T)} \tag{6.1}$$

$\tau(t, T)$  est la **fraction d'année** entre les date  $t$  et  $T$ .

Le **taux zéro-coupon linéaire**  $L(t, T)$  prévalant entre la date  $t$  et la maturité  $T$  correspond au rendement annuel d'un investissement de  $P(t, T)$  d'unités monétaires à la date  $t$  permettant d'obtenir une unité monétaire à la date  $T$  tel que intérêts s'accroissent proportionnellement au temps :

$$P(t, T)(1 + L(t, T)\tau(t, T)) = 1$$

2.  $\epsilon_t$  sont i.i.d avec pour moyenne  $\mathbb{E}[\epsilon_t] = 0$  et variance  $\text{Var}(\epsilon_t) = 1$ .  $\epsilon_t$  ne sont pas nécessairement un bruit blanc fort gaussien.

3. Cf. l'article de Christensen et al. [10]

ou encore

$$P(t, T) = \frac{1}{(1 + \tau(t, T)L(t, T))} \quad (6.2)$$

Un **contrat forward** permet à son détenteur de recevoir à la date  $S$  (avec  $S > T > t$ ) un paiement **fixe** d'intérêts calculés sur la période  $[T, S]$  pour un capital  $K$  contre le paiement d'intérêt **variable** (le taux étant **linéaire**) calculés également sur la période  $[T, S]$  pour un même capital  $K$ .

La valeur de ce contrat en  $S$  est

$$K\tau(T, S)(F - L(T, S))$$

La valeur en  $t$  est <sup>4</sup>

$$K(\tau(T, S)FP(t, S) - \frac{1}{P(T, S)} \times P(T, S) \times P(t, T) - P(t, S))$$

i.e. le taux  $F$  est déterminé tel que la valeur du contrat forward est nulle en  $t$

$$F(t, T, S) := \frac{1}{\tau(S, T)} \left( \frac{P(t, T)}{P(t, S)} - 1 \right) \quad (6.3)$$

Le **taux forward instantané** à partir de la définition du contrat forward est défini par :

$$\begin{aligned} f(t, T) &= \lim_{S \rightarrow T^+} F(t, T, S) \\ &= - \frac{\partial \ln P(t, T)}{\partial T} \end{aligned} \quad (6.4)$$

Le taux forward instantané est donc le taux d'un contrat forward lorsque  $S$  tend vers  $T$  i.e.

$$f(t, T) \simeq F(t, T, T + \delta t)$$

Par la suite on notera

$$f_t(\tau) := f(t, t + \tau)$$

avec  $\tau := \tau(t, T) =: T - t$

En utilisant les relations (6.4) et (6.1) on a

$$R(t, T) := \frac{1}{\tau(t, T)} \int_t^T f(t, u) du \quad (6.5)$$

Par la suite on notera

$$R_t(\tau) := R(t, t + \tau)$$

Un **swap de taux d'intérêt standard** est un contrat dans lequel deux contreparties s'engagent mutuellement à se verser des flux d'intérêt basés sur un capital pendant une durée déterminée suivant des fréquences différentes et une base calendaire contractuelle. L'une des contrepartie paie des flux **fixes**, il s'agit de la **jambe fixe**, l'autre partie paie des flux **variables** basés sur un indice de taux variable, il s'agit de la **jambe variable**.

On note  $\mathcal{T} := \{T_{\alpha+1}, \dots, T_\gamma\}$  l'ensemble des dates de paiements de la contrepartie *jambe fixe* et  $\mathcal{T}' := \{T_{\alpha'+1}, \dots, T_\gamma\}$  celles de la contrepartie *jambe variable*. La valeur du contrat est déterminée en valorisant les flux des jambes à la date  $t$  pour tout  $t \leq T_\alpha$  :

---

4. En utilisant la définition (6.1.1) page 103



- $C \times K \times \sum_{i=\alpha+1}^{\gamma} \tau(T_{i-1}, T_i)$
- $K \times \sum_{i=\alpha'+1}^{\gamma} \tau(T_{i-1}, T_i) L(T_{i-1}, T_i)$

Le taux de swap est le taux  $S_{\alpha, \gamma}(t) := C$  tel que le contrat soit *fair* i.e.

$$S_{\alpha, \gamma}(t) \sum_{i=\alpha+1}^{\gamma} \tau(T_{i-1}, T_i) P(t, T_i) = P(t, T_{\alpha}) - P(t, T_{\gamma})$$

$$S_{\alpha, \gamma}(t) := \frac{P(t, T_{\alpha}) - P(t, T_{\gamma})}{\sum_{i=\alpha+1}^{\gamma} \tau(T_{i-1}, T_i) P(t, T_i)} \quad (6.6)$$

### 6.1.2 Fonctionnelle de Nelson-Siegel

Nelson-Siegel[28] ont proposé un **paramétrisation** de taux forward instantané à quatre paramètres, trois paramètres de forme  $(\beta_0, \beta_1, \beta_2)$  et un paramètre d'échelle  $\lambda$  :

$$f_t(\tau) = \beta_0 + \beta_1 e^{-\frac{\tau}{\lambda}} + \beta_2 \frac{\tau}{\lambda} e^{-\frac{\tau}{\lambda}}$$

Cette paramétrisation introduit trois **facteurs** : une constante  $f_0$ , une fonction à **décroissance exponentielle**  $f_1$  et un **polynôme de Laguerre**<sup>5</sup>  $f_2$

$$\begin{cases} f_0 := 1 \\ f_1 := e^{-\frac{\tau}{\lambda}} \\ f_2 := \frac{\tau}{\lambda} e^{-\frac{\tau}{\lambda}} \end{cases}$$

Elle permet d'obtenir différentes formes de courbe de taux d'intérêt. En effet

- la constante représente le niveau de taux long.
- la fonction à décroissance exponentielle permet d'obtenir des courbes croissantes ou inversées en fonction du signe du paramètre  $\beta_1$  ( $\beta_1 < 0$  courbe inversées).
- la fonction de Laguerre permet d'obtenir de contrôler la convexité en fonction du signe de  $\beta_2$ .

En utilisant le résultat (6.5) on obtient :

$$R_t(\tau) = \beta_0 + (\beta_1 + \beta_2) \left( \frac{1 - e^{-\frac{\tau}{\lambda}}}{\frac{\tau}{\lambda}} \right) + \beta_2 e^{-\frac{\tau}{\lambda}} \quad (6.7)$$

#### Hypothèse 6.1.1

Nous allons supposer maintenant que les paramètres de la fonctionnelle de Nelson-Siegel  $(\lambda, \beta_0, \beta_1, \beta_2)$  dépendent du **temps** i.e.

$$R_t(\tau) = \beta_{0t} + (\beta_{1t} + \beta_{2t}) \left( \frac{1 - e^{-\frac{\tau}{\lambda_t}}}{\frac{\tau}{\lambda_t}} \right) + \beta_{2t} e^{-\frac{\tau}{\lambda_t}} \quad (6.8)$$

---

5. Polynôme de Laguerre : produit d'un polynôme et une exponentielle.

### 6.1.3 Estimation

On souhaite déterminer les paramètres de la fonctionnelle de Nelson-Siegel pour chaque date d'observation historique. Pour cela on doit résoudre le problème de moindre carré suivant en utilisant<sup>6</sup> l'expression des taux de swaps en fonction des paramètres :

$$\min_{\beta_{0t}, \beta_{1t}, \beta_{2t}, \lambda_t} \sum_{T \in T_1, \dots, T_N} (S_{\alpha, \gamma}(t, T) - S_{\alpha, \gamma}^{obs}(t, T))^2 \quad (6.9)$$

où les  $T_i$  sont les maturités des taux de swaps observés sur le marché.

Les caractéristiques de la fonctionnelle de Nelson Siegel rendent l'optimisation complexe. Nous allons utiliser deux algorithmes

1. un algorithme à évolution différentielle permettant de déterminer des conditions initiales
2. un algorithme à directions de descente permettant de déterminer la solution globale à partir des conditions initiales initialement déterminées.

#### 6.1.3.1 Paramètres de Nelson-Siegel pour la devise EUR

Sur la période du 20/10/1999 au 11/09/2014 nous allons étudier l'évolution des paramètres de Nelson-Siegel estimés sur les données EUR.

$$\min_{\beta_{0t}, \beta_{1t}, \beta_{2t}, \lambda_t} \sum_{T \in T_1, \dots, T_N} (S_{\alpha, \gamma}(t, T) - S_{\alpha, \gamma}^{obs}(t, T))^2 \quad (6.10)$$

#### Proposition 6.1.1 (Résolution du programme d'optimisation)

On pose  $x_t := (\beta_{0t}, \beta_{1t}, \beta_{2t}, \lambda_t)$ .

Pour résoudre le problème (6.10), nous allons effectuer les étapes suivantes :

- POUR DÉTERMINER LE POINT INITIAL  $x_{t-H} \in \mathbb{R}^4$  ON UTILISE L'ALGORITHME À ÉVOLUTION DIFFÉRENTIELLE
- ON APPLIQUE L'ALGORITHME BFGS EN UTILISANT CHAQUE FOIS LA SOLUTION PRÉCÉDENTE  $x_{t-i}^*$  COMME POINT INITIAL.
- SI  $|x_{t-i+1}^* - x_{t-i}^*|$  ON UTILISE DE NOUVEAU L'ALGORITHME À ÉVOLUTION DIFFÉRENTIELLE.

L'idée du **couplage** de ces deux algorithmes est d'accélérer les temps de calculs. Le choix de la condition initiale est délicat pour l'algorithme de descente. D'autre part, les paramètres d'une date à une autre ne devrait pas évoluer drastiquement.

Le graphique suivant présente l'évolution du paramètre  $\lambda$  estimé sur la période 20/10/1999 au 11/09/2014 sur la courbe<sup>7</sup> de swap de taux d'intérêt (1Y, ..., 10Y, 15Y, 20Y) :

6. En utilisant les relations (6.1) page 103 et (6.8) page 105

7. Source : Bloomberg

### Série temporelle du paramètre lambda

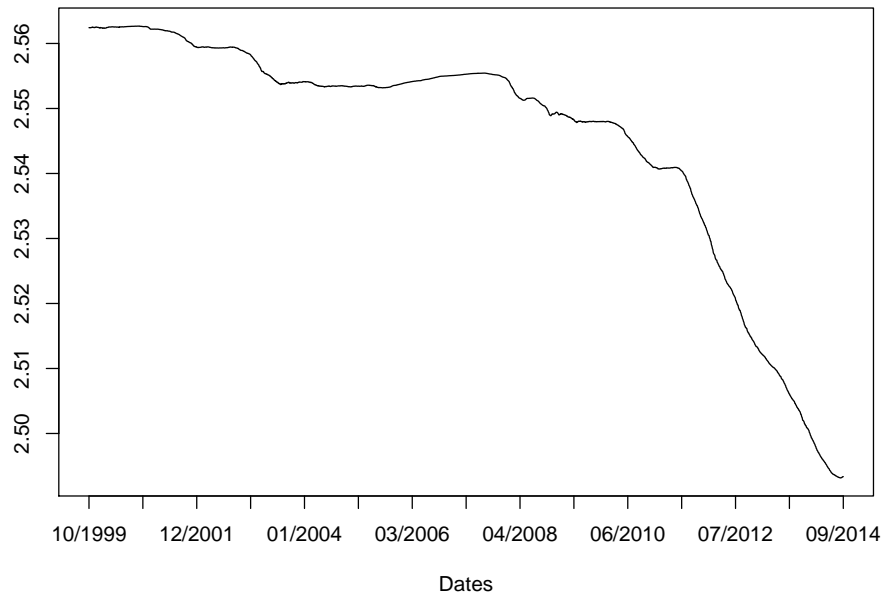


FIGURE 6.1 – Paramètre lambda estimé sur la période 20-10-1999  $\mapsto$  09-11-2014

Expliquer pourquoi c'est peu.

On peut remarquer que le paramètre  $\lambda$  évolue **peu** au cours du temps même s'il semble avoir un léger décrochage pendant la crise de la dette souveraine dans de la zone euro :

Sur la période historique

- Le minimum est : 2.493166
- Le maximum est : 2.562675
- La moyenne est : 2.545487
- L'écart-type est de 1,839454%

Il **n'est** donc **pas** indispensable de l'**estimer**. C'est un paramètre d'échelle qui n'est pas sensé évoluer sensiblement au cours du temps.

D'autre part, le paramètre  $\lambda$  **fixé** rend le programme d'optimisation plus simple.

### Hypothèse 6.1.2

Par la suite nous allons choisir  $\lambda_t = \lambda = 2.5 \quad \forall t$ .

*Dans la littérature, on exprime souvent  $\lambda$  à partir de maturités exprimées en nombre de mois plutôt qu'en nombre de mois.*

Les graphes ci-dessous montrent l'évolution des paramètres de formes de Nelson-Siegel et leur interprétation suggérée par Diebold et Lie [17] :

— Le facteur de forme  $\beta_0$  est un **facteur de niveau** représentant le niveau de taux zéro-coupon **10 ans**.  
**Expliquer le choix du 10 ans**

Ci-dessous le graphe des séries de taux zéro-coupon (en noir) et  $\beta_0$  (en rouge) sur la période 1-1-1999 → 03-03-2014.

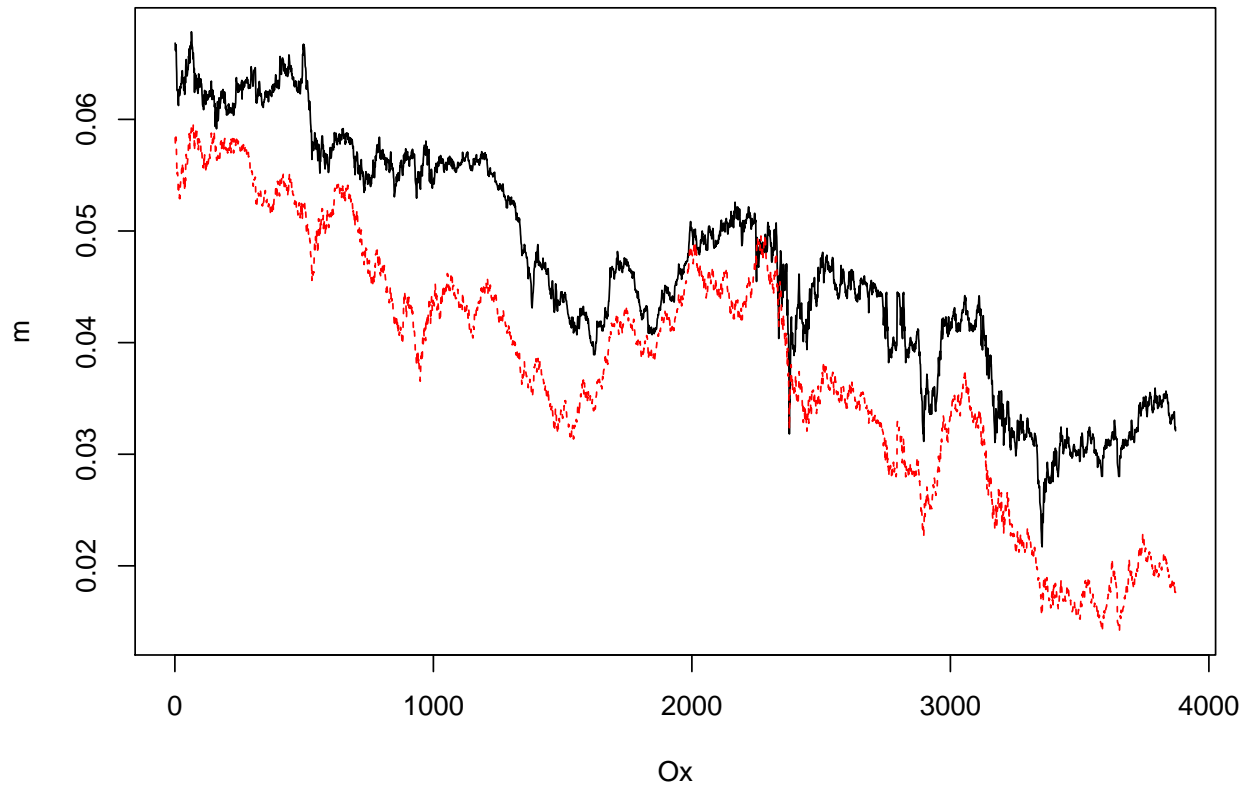


FIGURE 6.2 – Séries historiques des  $\beta_0$  et du taux ZC 10Y - Calcul via la forme estimée de Nelson-Siegel sur la période 1-1-1999 → 03-03-2014

Le corrélation entre les deux séries est élevé elle est de 0.9489536.

- Le facteur de forme  $\beta_1$  est un facteur de forme représentant l'opposé du spread entre le taux zéro-coupon 10Y et le taux zéro-coupon 3M

Ci-dessous le graphe des séries de l'opposé du spread (en noir) et  $\beta_1$  (en rouge) sur la période 1-1-1999 → 03-03-2014.

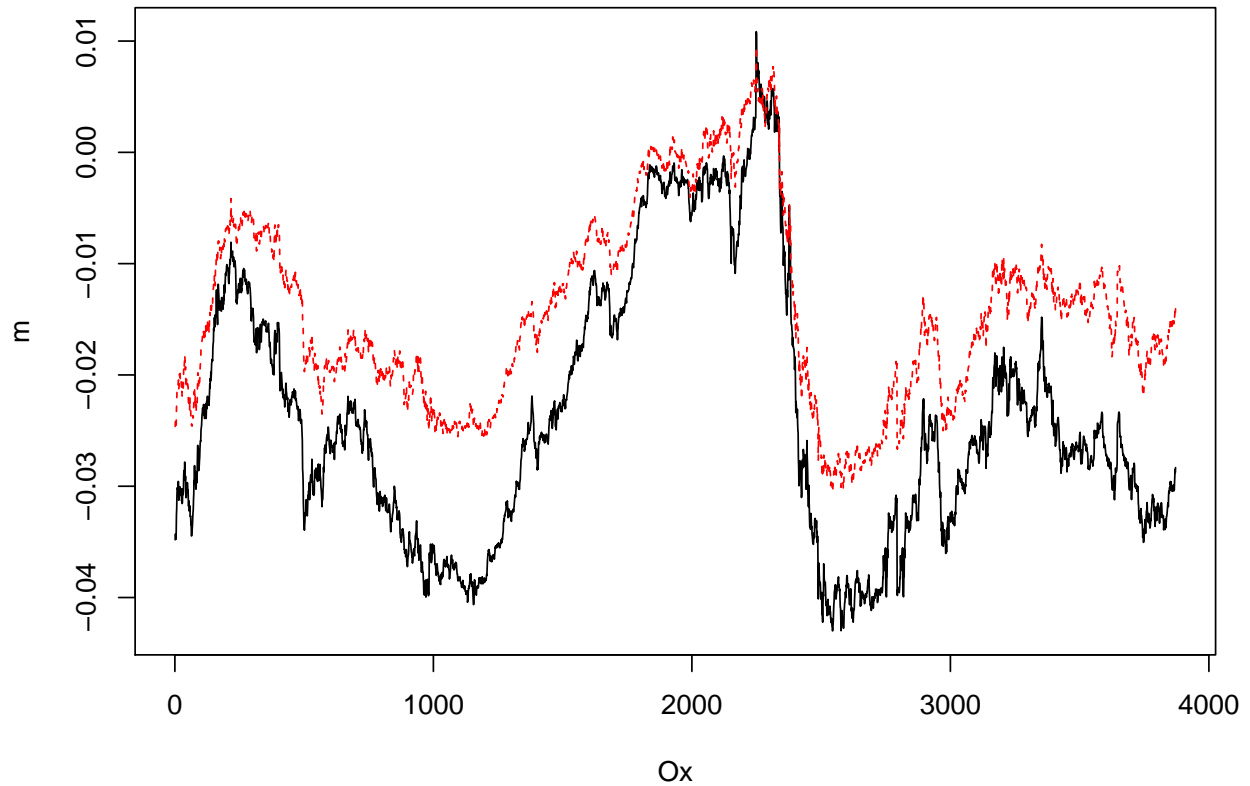


FIGURE 6.3 – Paramètre lambda estimé sur la période 20-10-1999 ↦ 09-11-2014

La corrélation entre les deux séries est de 0.9752922.

D'autre part, intrinsèquement le modèle de Nelson-Siegel indique que

$$\beta_0 + \beta_1 = \lim_{\tau \rightarrow 0} y_t(\tau)$$

Ci-dessous le graphe des séries du taux court et de  $\beta_0 + \beta_1$  (en rouge) sur la période 1-1-1999  $\rightarrow$  03-03-2014.

### Séries temporelles du facteur court terme

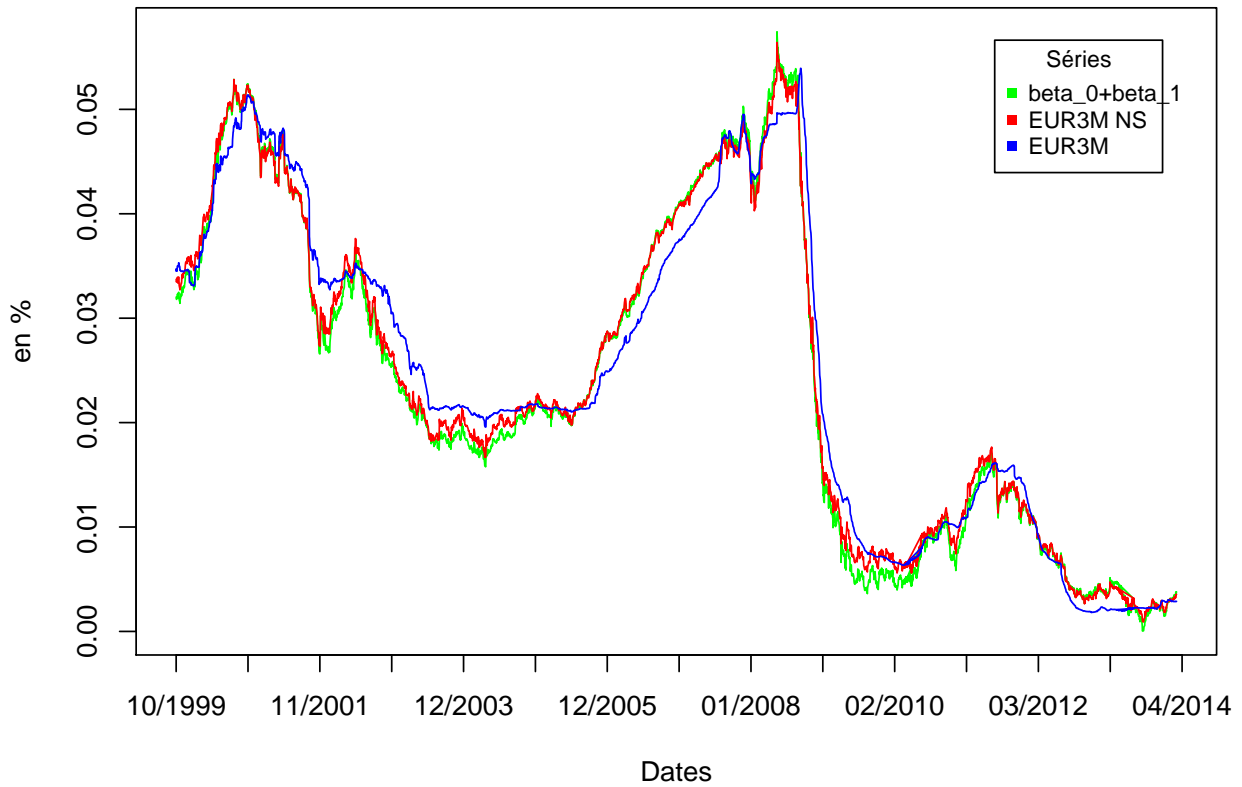


FIGURE 6.4 – Séries historiques des  $\beta_0 + \beta_1$  et du taux zéro-coupon 3M- Calcul via la forme estimée de Nelson-Siegel sur la période 1-1-1999  $\rightarrow$  03-03-2014

La corrélation entre les séries  $\beta_0 + \beta_1$  est de 0.9990092.

#### 6.1.4 Modélisation AR(1) des paramètres $\beta_i$

##### Hypothèse 6.1.3 (Hypothèses sur les paramètres $\beta$ de Nelson-Siegel)

On considère que chaque paramètre  $\beta_i$  est modélisé par un modèle **autorégressif** AR(1) :

$$\beta_{it} = \alpha_i + \phi_i \beta_{it-1} + \epsilon_{it}$$

Les graphiques suivants des ACF et des PACF des séries  $\beta_i$  montrent

- une décroissance exponentielle faible des fonctions d'auto-corrélations
- une fonction d'auto-corrélation partielle significativement nulle

Ces premiers résultats permettent de considérer le modèle  $AR(1)$  pour les séries  $\beta_i$ .

L'estimation des processus AR(1) donne pour

$$\beta_0 : \beta_{0t} = 0.04694398 + 0.99918310\beta_{0t-1} + \widehat{\epsilon}_{0t}$$

$$\beta_1 : \beta_{1t} = -0.02355585 + 0.99857416\beta_{1t-1} + \widehat{\epsilon}_{1t}$$

$$\beta_2 : \beta_{2t} = -0.01398015 + 0.99285653\beta_{2t-1} + \widehat{\epsilon}_{2t}$$

On remarque que l'estimation des processus  $\beta_i$  indique que les processus sont proches de la **non-stationnarité** i.e. les  $\phi_i \simeq 1$ .

**Estimation de la constante**



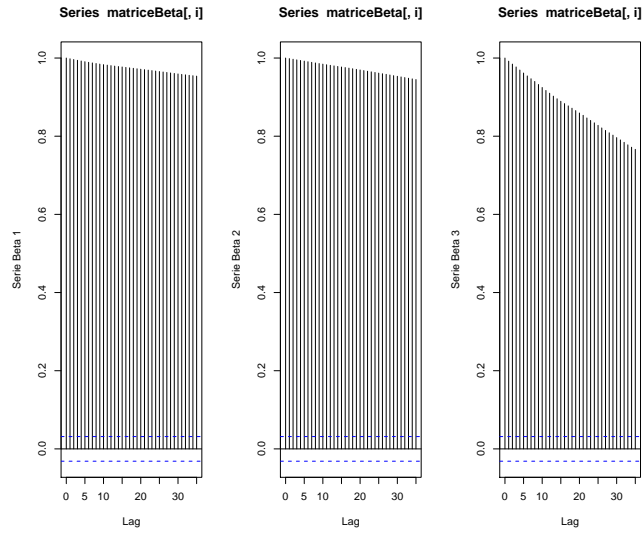


FIGURE 6.5 – Fonctions d'autocorrélation des séries  $\beta_i$

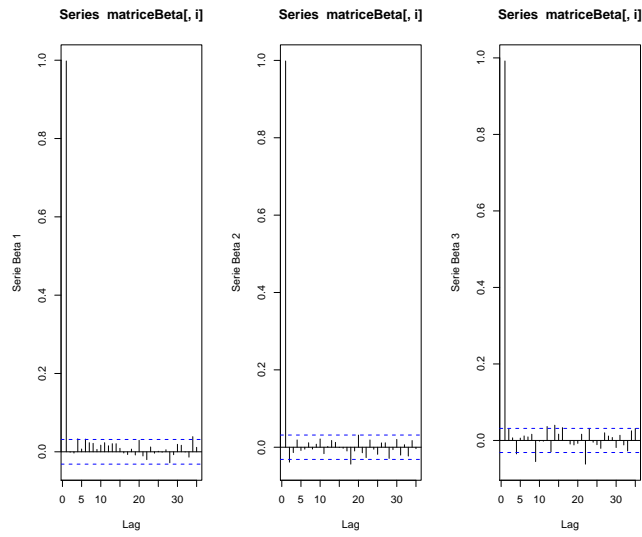


FIGURE 6.6 – Fonctions d'autocorrélation partielle des séries  $\beta_i$

Nous allons maintenant juger de l'adéquation du modèle AR(1) aux séries temporelles. Nous allons étudier les résidus.

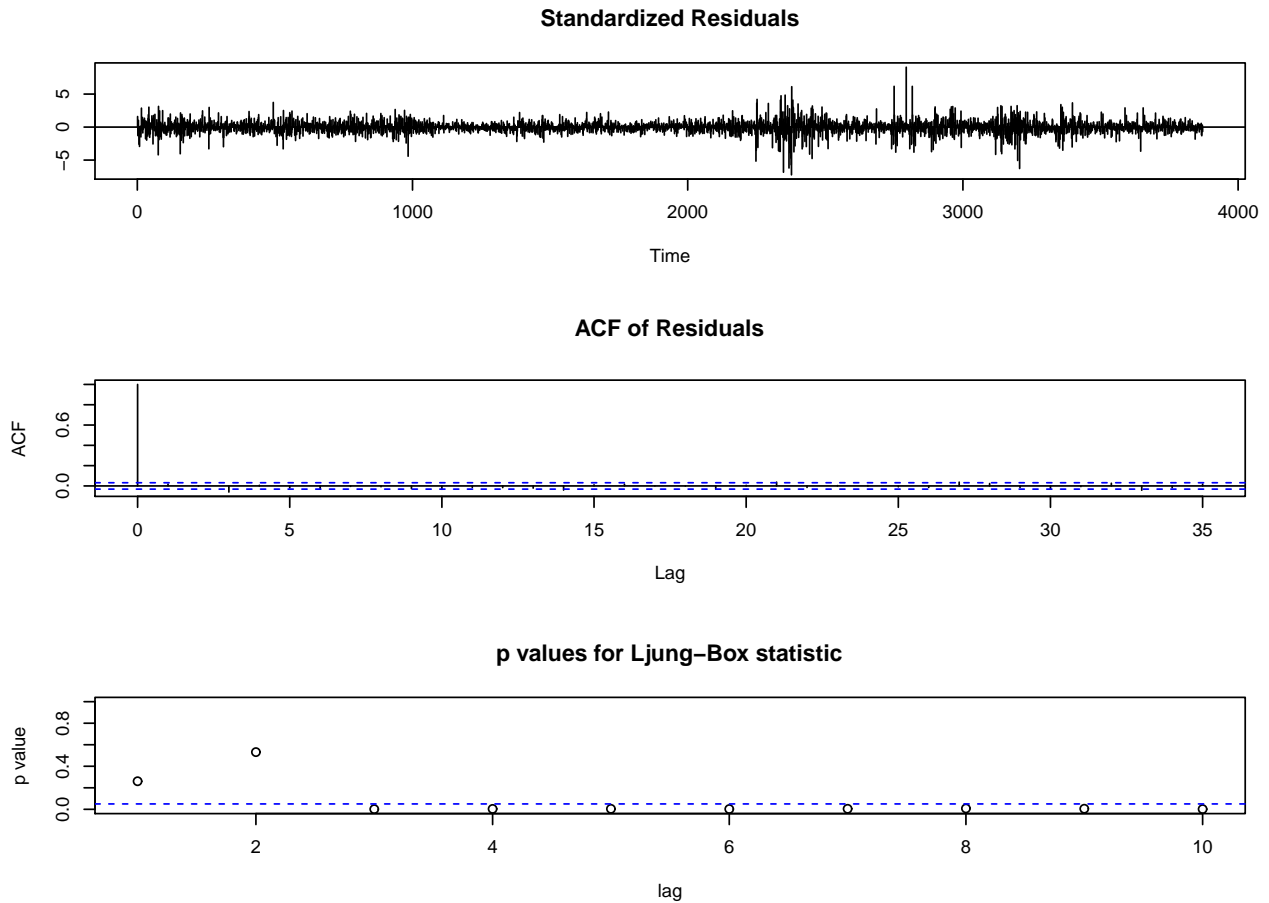


FIGURE 6.7 – Test d'indépendance des résidus

1. Le premier graphique représente la série des résidus.
2. Le deuxième graphique représente l'auto-corrélogramme des résidus.  
Nous observons que la fonction d'auto-corrélation vaut 1 pour  $h = 0$  et est significativement nulle pour  $h > 0$ .
3. Le troisième graphique représente le test d'indépendance de Jung-Box des résidus. Pour que le modèle soit valide les résidus estimés doivent suivre un bruit blanc fort en particulier les  $\epsilon_t$  et  $\epsilon_{t-k}$  qui doivent

être indépendants.

Nous observons que le test d'indépendance est **rejeté** en effet certaines p-value sont significativement **nulle**.

L'estimation a été effectuée sur la période 01/01/1999  $\mapsto$  03/03/2014.

Considérons les sous-périodes suivantes :

1. période avant crise sub-prime 01/01/1999  $\mapsto$  01/07/2007
2. période crise de sub-prime et interbancaire 01/07/2007  $\mapsto$  01/10/2009
3. période crise souveraine 01/10/2009  $\mapsto$  30/09/2012
4. période post-crise souveraine 01/10/2012  $\mapsto$  03/03/2014

Les graphes ci-dessous présente le test d'indépendance des résidus du modèle AR(1) sur chacune des périodes précitées.

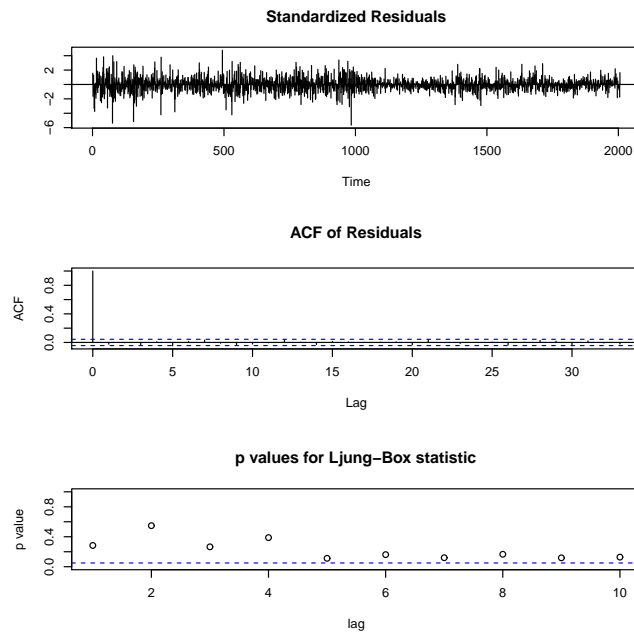


FIGURE 6.8 – Test d'indépendance des résidus pour la période avant crise

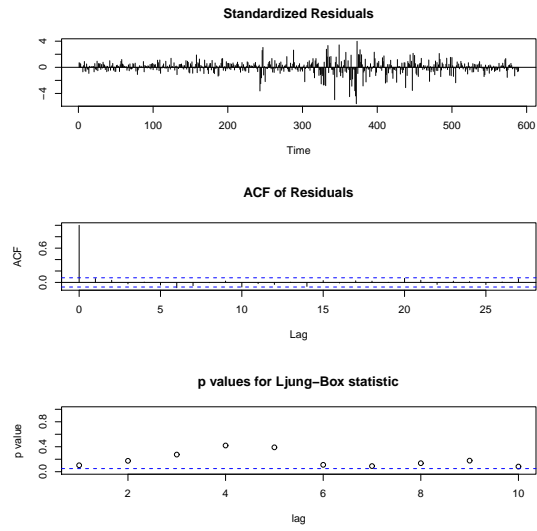


FIGURE 6.9 – Test d'indépendance des résidus pour la période crise sub-prime

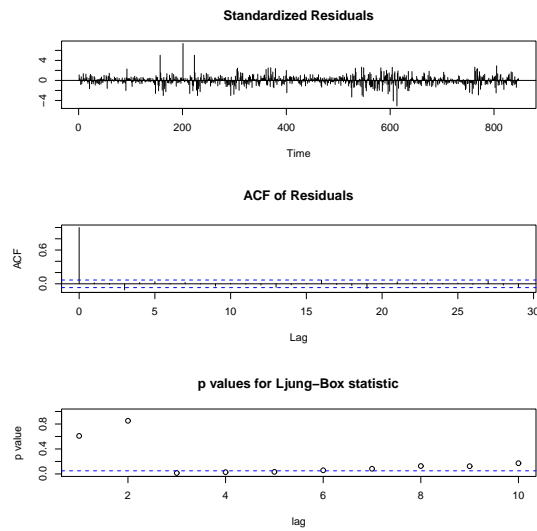


FIGURE 6.10 – Test d'indépendance des résidus pour la période crise souveraine

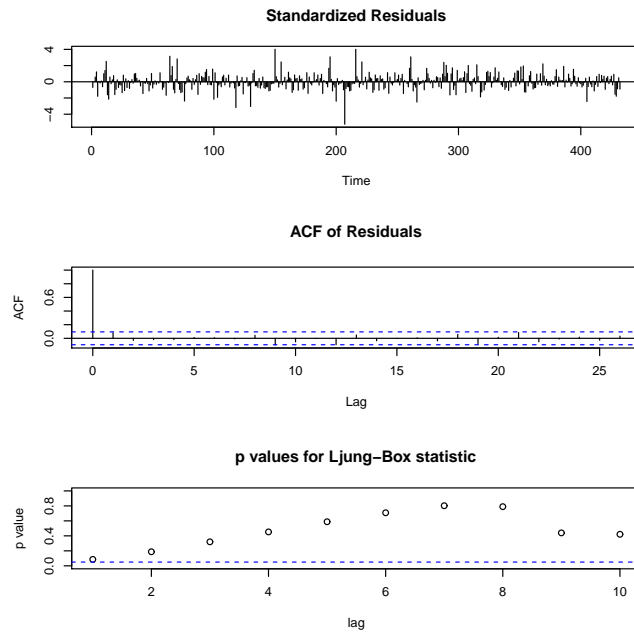


FIGURE 6.11 – Test d’indépendance des résidus pour la période d’après crise souveraine

Pour chaque **sous-période** le modèle AR(1) est **accepté**.

## 6.2 Modèle dynamique pour les actions

La poche "action" de l'actif des entreprises d'assurance, même si elle représente une poche beaucoup plus petite<sup>8</sup> que la poche obligataire, constitue un risque **important**. En effet, la poche action présente une forte volatilité de sa valeur économique et qui est illustré dans le  $SCR_{\text{Marché}}$  souvent élevé.

En pratique, on choisit un indice action comme *proxy* de la poche "action".

### 6.2.1 Analyse descriptive

Dans ce mémoire, nous choisissons l'indice CAC40. Mais ce choix doit être effectué en adéquation avec la composante diversification (principalement R332-20) du portefeuille d'actif de l'entreprise d'assurance.

Ci-dessous la série des **log**-valeurs de l'indice CAC40<sup>9</sup> :

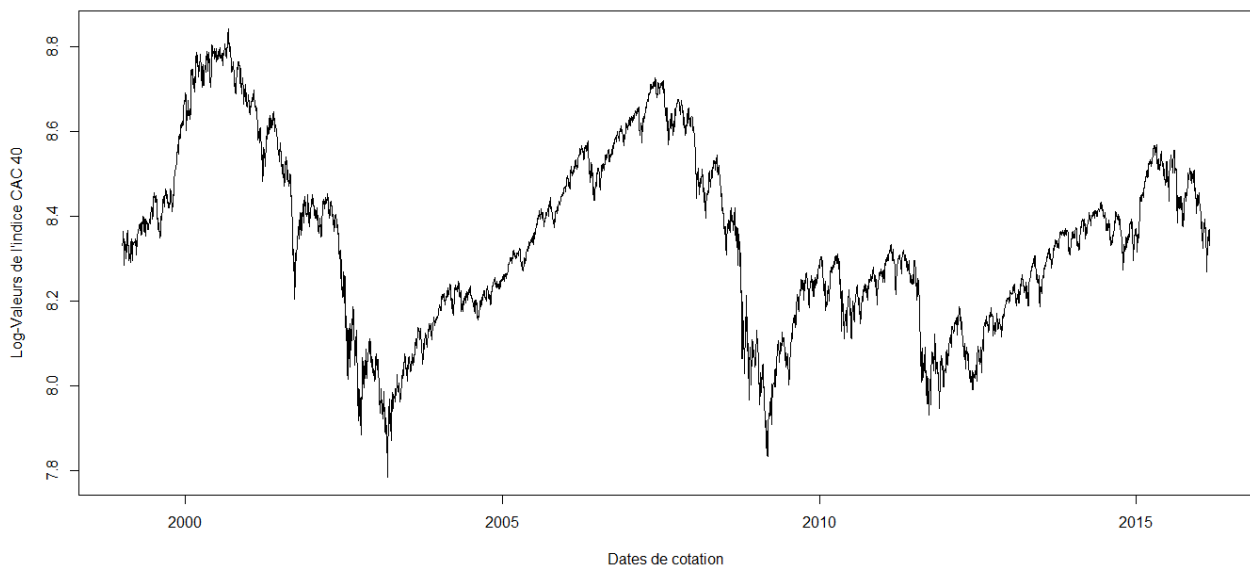


FIGURE 6.12 – Evolution du CAC40 sur la période 04-01-1999  $\mapsto$  26-02-2016

---

8. Fin 2013, la poche action des entreprises d'assurance ayant réalisé l'exercice préparatoire SII représente environ 6,3% de l'actif économique Cf. le document [1].

9. Source Yahoo valeurs de clôture

La dynamique de l'indice ne semble pas montrer une indépendance des moments par rapport au temps. Cela suggère une non-stationnarité du processus indicelle.

Le test de stationnarité de Dickey-Fuller nous donne une *p-value* calculée très supérieure au seuil de signification de 0,05%, nous ne pouvons donc pas rejeter l'hypothèse nulle de non-stationnarité.

Ainsi, on s'intéresse souvent plutôt à la série **différenciée** des **log-rendements**<sup>10</sup> journaliers :

$$x_t = \ln \frac{s_t}{s_{t-1}}$$

où  $s_t$  est la valeur de l'indice boursier observée à la date  $t$  et  $s_{t-1}$  la valeur du (jour) ouvré précédant.

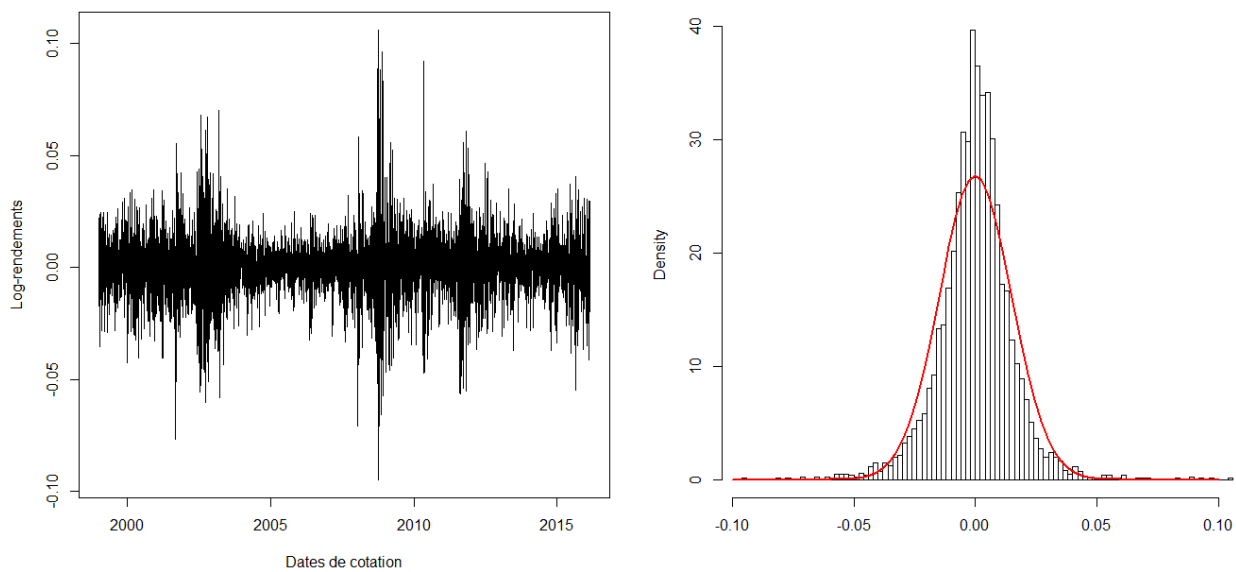


FIGURE 6.13 – Log-rendements du CAC 40 sur la période 04-01-1999  $\mapsto$  26-02-2016

Le tableau suivant récapitule les statistiques usuelles :

Min	1er Qu.	Médiane	Moyenne	3er Qu.	Max	Skewness	Kurtosis
-9,472e-02	-7,319e-03	3,773e-04	9,010e-06	7,798e-03	1,059e-01	-0.008200641	7.461133

10. En utilisant  $\ln(1+x) \sim x$ , on retrouve la mesure de variation relative.

Pour modéliser l'évolution des indices "action", on utilise souvent le modèle populaire de Black-Sholes<sup>11</sup> (BS) en particulier en risque-neutre. La discrétisation du log-processus est donnée par le schéma d'Euler<sup>12</sup> :

$$S_t = S_{t-1} \exp \left\{ \left( r - \frac{\sigma_S^2}{2} \right) + \sigma_S \epsilon_t \right\}$$

La modélisation BS suppose

- une distribution **gaussienne** du log-processus - loi du mouvement brownien standard
- une volatilité **constante**.

Plusieurs faits stylisés<sup>13</sup> contredisent ces hypothèses. En effet, on observe souvent pour les indices boursiers et en particulier pour le CAC 40 :

- la distribution empirique est à queues **épaisses**<sup>4</sup>. La *skewness* et la valeur (élevée) du *kurtosis* ne correspondent pas aux valeurs de référence de la loi normale. Si nous effectuons un test de Jarque-Bera, la valeur de la p-value est élevée ce qui confirme le **rejet** de l'hypothèse de **normalité**.
- la variance de la distribution empirique **ne** semble **pas** être **stable** dans le **temps**. On observe en particulier des **regroupements** d'extrêmes c'est ce qu'on appelle **volatility clustering** : des périodes de fortes volatilités succèdent des période de faibles volatilités, et inversement. Utiliser une volatilité (constante) **historique** conduit à **sous** ou **sur**-estimer les prix risque-neutre des options.

Ainsi plusieurs modèles ont été proposés pour prendre en compte ces phénomènes empiriques :

- à **sauts**.

On considère que l'actif suit un processus BS **entre** deux temps de **saut**  $\tau_i$  d'un processus de Poisson  $N_t$  et vérifie **en**  $\tau_i$

$$S_{\tau_i} = S_{\tau_i^-} (1 + U_i)$$

On a l'expression générale :

$$S_{t+1} = S_t \left( \prod_{i=N_t}^{N_{t+1}} (1 + U_i) \right) \exp \left\{ \left( r - \frac{\sigma_S^2}{2} \right) + \sigma_S \epsilon_t \right\}$$

$$\ln \frac{S_{t+1}}{S_t} = \sum_{i=N_t}^{N_{t+1}} \ln(1 + U_i) + \left( r - \frac{\sigma_S^2}{2} \right) + \sigma_S \epsilon_t$$

On suppose souvent que  $1 + U_i$  sont des variables aléatoires de lois **log-normales**. Ainsi, on a des rendements gaussiens conditionnellement à  $N_t$ .

---

11. Brownien géométrique avec volatilité constante voire déterministe du temps.

12. Appelé aussi schéma de discrétisation d'Euler-Maruyama. La discrétisation de Milstein du log-processus est identique à celle d'Euler.

13. Propriétés statistiques communes - Stylised facts. Cf. travaux de Mandelbrot



On retrouve les modélisations à Poisson-composée de Lunberg ou le modèle discret à sauts de Merton<sup>14</sup>.  
— à changement de régimes (markovien).

Par exemple dans le GSE d'Ahlgrim, les indices actions sont modélisés via un modèle à régime de type Hardy.

— à volatilité stochastique.

Ils constituent une alternative, en particulier en risque-neutre<sup>15</sup>, au modèle de BS :

$$\begin{aligned}\frac{dS_t}{S_t} &= r_t S_t + \sigma_t dW_t \\ d\sigma_t &= \alpha_t dt + \tilde{\sigma} dB_t\end{aligned}$$

— à volatilité conditionnelle déterministe de type ARCH :

$$\epsilon_t = \sqrt{h_t} u_t$$

avec  $h_t$  est une fonction **déterministe** du passé.

Remarquons qu'ils diffèrent par rapport aux modèles à volatilité stochastique par le fait que dans leur version discrète, on a  $\epsilon_t = \sqrt{h_t} u_t$  avec  $h_t \neq \mathbb{E}_{t-1}[\epsilon_t^2]$

Dans ce mémoire, nous choisissons de modéliser les séries "action" par des modèles **auto-régressifs** avec une structure **hétéroscédastique**. Ils constituent une **alternative** discrète au modèle de BS :

## 6.2.2 Modélisation ARIMA – GARCH

Pour déterminer le modèle le plus adéquat de type ARIMA – GARCH, nous allons utiliser

- la procédure d'identification de Box et Jenkins pour identifier les paramètres  $(p, d, q)$  du modèle ARIMA
- le test d'Engle pour identifier une structure ARCH

---

14. Cf. l'article de F. Planchet et PE. Thérond [32].

15. Cf. le modèle de Heston.

Ci-dessous les auto-corrélogrammes et auto-corrélogrammes partiels pour des log-séries et des log-séries différenciées.

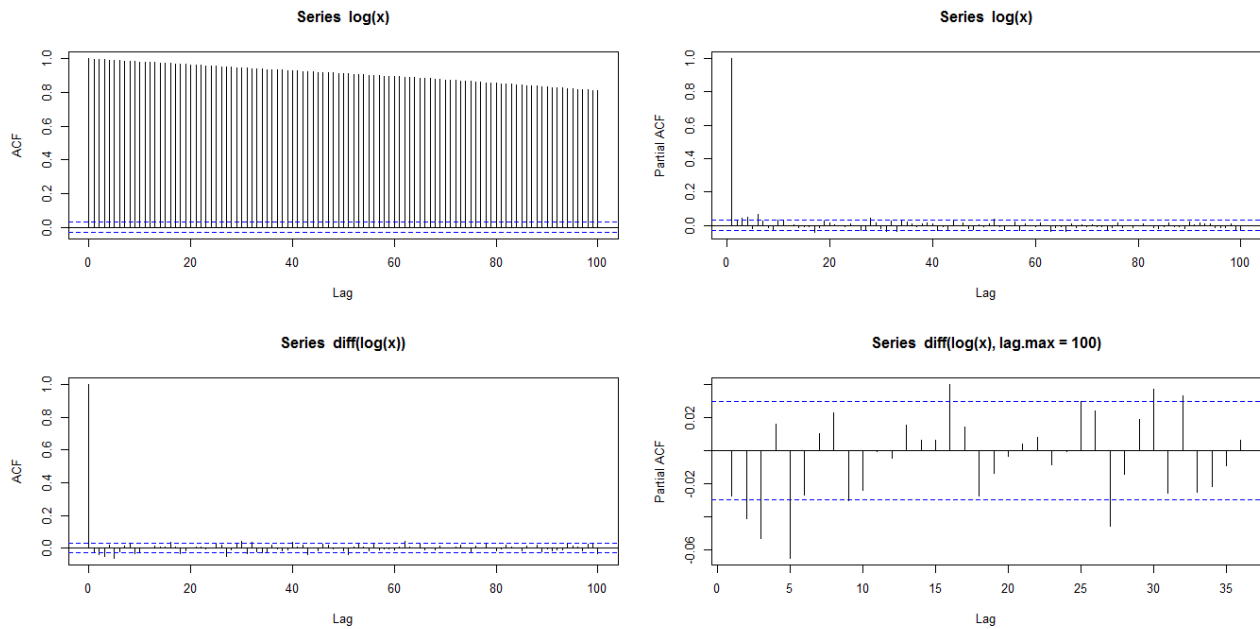


FIGURE 6.14 – Log-rendement du CAC40 sur la période 04-01-1999  $\mapsto$  26-02-2016

Nous remarquons :

- le graphique en haut à gauche représente l’auto-corrélogramme de la **log**-série (de l’indice CAC 40). La décroissance lente incite à différencier la série.
- le graphique en haut à droite représente l’auto-corrélogramme partiel de la log-série. La nullité de valeurs au-delà de l’ordre 1 suggère un modèle de type ARIMA(1,0,0) i.e de type AR(1)
- Le graphique en bas à gauche représente l’auto-corrélogramme de la log-série **différenciée**. On constate la nullité des coefficients pour tout  $h > 0$  : **non-corrélation des rendements**
- Le graphique en bas à droite représente l’auto-corrélogramme partiel de la log-série différenciée. Tous les coefficients sont significativement nuls.
- Les deux derniers points montre un comportement de type bruit blanc pour la série différenciée et donc de type ARIMA(0,1,0) pour la log-série.

Cette analyse nous suggère donc une modélisation plus complexe de type ARIMA( $p, d, q$ ). Nous avons indiqué

que dans le paragraphe sur les critères de validation<sup>16</sup>, la parcimonie est un critère que nous recherchons, nous allons donc comparer des modèles ARIMA( $p, d, q$ ) suivants avec des ordres **inférieurs** à **deux** :

Modèle ARIMA	Statistique AIC
ARIMA(1,1,0)	-24 458
ARIMA(1,1,1)	-24 476.58
ARIMA(1,1,2)	-24 476.86
ARIMA(2,1,0)	-24 464
ARIMA(2,1,1)	-24 476.92
ARIMA(2,1,2)	-24 474

Le critère AIC nous suggère de choisir celui qui a la statistique AIC le plus faible c'est-à-dire le modèle ARIMA(2, 1, 1) pour le processus  $\{\ln X_t\}_{t \in \mathbb{Z}}$ .

Nous vérifions bien que les résidus sont un **bruit blanc** faible : moyenne nulle, auto-corrélations totales et partielles nulles :

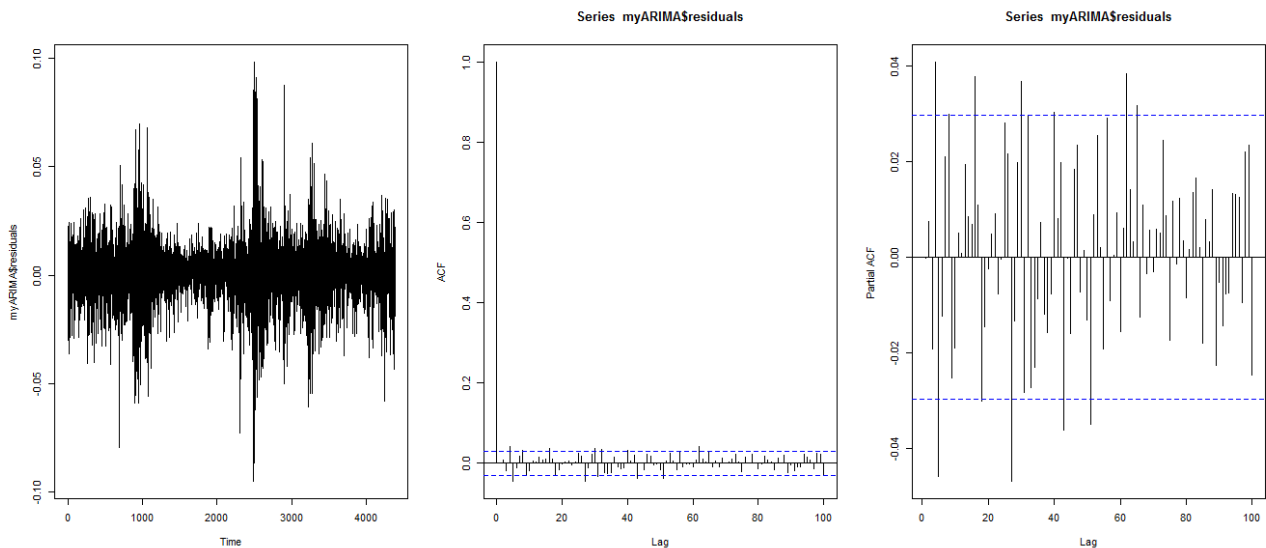


FIGURE 6.15 – Résidus du modèle ARIMA(2, 1, 1) sur la période 04-01-1999  $\mapsto$  26-02-2016

Nous effectuons les tests suivants

- L'hypothèse d'auto-corrélations nulles, n'est **pas** rejetée par le test de **Box-Pierce** : probabilité critique **grande** pour tout lag.

---

16. Cf. page

- L'hypothèse nulle de non-stationnarité est rejetée par le test de Dickey-Fuller augmentée (pValue petite)
- L'hypothèse nulle de normalité est rejetée par le test de Jarque-Bera (pValue petite)

D'autre part, nous observons sur le graphique de gauche des résidus un phénomène de *volatility clustering*. Nous cherchons donc à tester l'hypothèse de structure ARCH dans les résidus.

Nous observons dans le graphique ci-dessous que l'hypothèse d'auto-corrélations nulles pour différents *lags* des **carrés** des résidus n'est pas être vérifiée

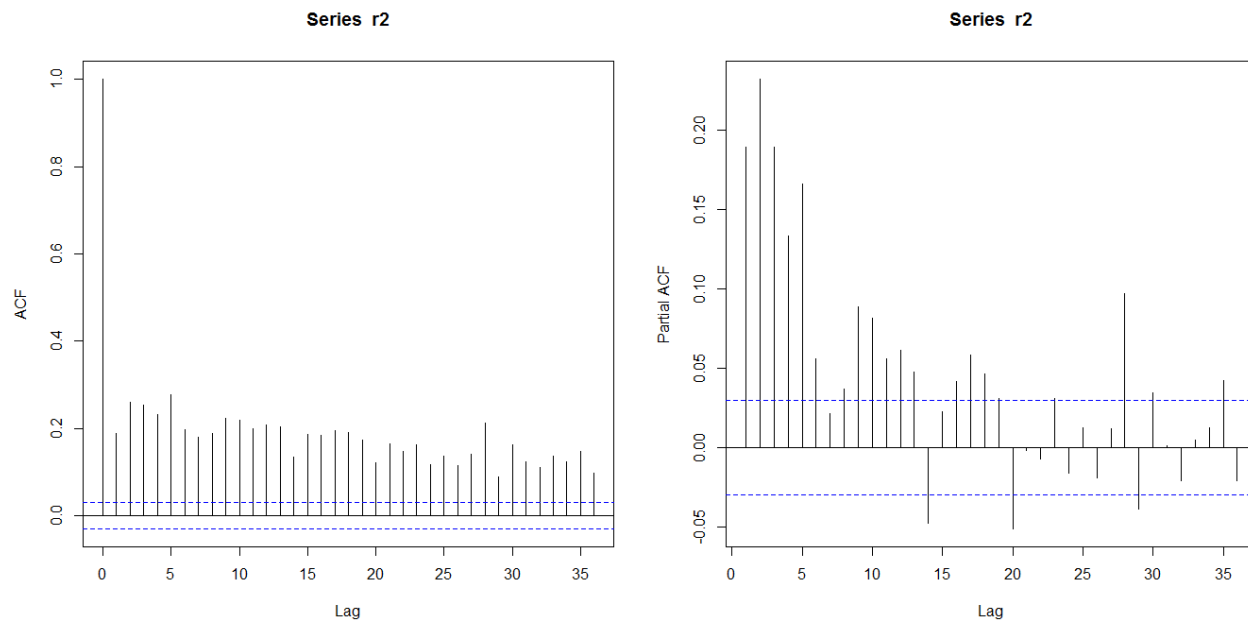


FIGURE 6.16 – Auto-corrélogramme total et partiel pour le carré des résidus du modèle  $ARIMA(2, 1, 1)$  sur la période 04-01-1999  $\mapsto$  26-02-2016

Ceci est confirmé par le test de Box-Pierce qui présente une probabilité critique petite pour les auto-corrélations des carrés des résidus.

Mais testons plus spécifiquement la structure ARCH via le test d'Engle. Rappelons que le test d'Engle est basé sur la corrélation des carrés des résidus. La probabilité critique est grande l'hypothèse alternative de structure ARCH est acceptée<sup>17</sup>.

Nous devons maintenant déterminer la structure ARCH pour les résidus du modèle  $ARIMA(2, 1, 1)$ .

17. Nous utilisons la fonction ARCHTest du package "FinTS" de R Cf. [78].

Nous testons les modèles suivants :

- GARCH(0, 1) i.e. ARCH(1)
- GARCH(1, 1)

Nous utilise de nouveau le critère AIC pour départager les modèles.

Modèle GARCH	Statistique AIC
GARCH(0,1)	-24 737
GARCH(1,1)	-25 823

Nous choisissons donc le modèle ARIMA(2, 1, 1) – GARCH(1, 1) pour le log-processus  $\{\ln S_t\}_{t \in \mathbb{Z}}$  i.e. le processus  $X_t$  défini par  $(1 - L) \ln S_t$  est un processus ARMA(2, 1) – GARCH(1, 1)

$$\left\{ \begin{array}{l} X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t \\ \epsilon_t = h_t u_t \\ h_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}^2 \end{array} \right.$$

Nous vérifions que les résidus **standardisés**

$$\forall t \in [t - H + 1, t - 1] \quad u_t = \frac{X_t - \hat{\phi} X_{t-1}}{\hat{h}_t} \quad \text{avec} \quad \hat{h}_t = \hat{\alpha}_0 + \hat{\alpha}_1 \epsilon_{t-1} + \hat{\beta}_1 \hat{h}_{t-1}$$

sont un bruit blanc **gaussien**.

On effectue les tests suivants :

- La distribution des résidus n'est pas normale, en effet l'hypothèse de **normalité** est **rejetée** par le Jarque-Bera. La p-value est nulle.
- Les résidus ne sont pas **auto-corrélés**, en effet le test de Jung-Box est **rejeté**, la p-value est grande.

## 6.3 Modèle dynamique pour le crédit

Le **crédit** est souvent peu modélisé, mais depuis la crise de 2008 et du fait que le portefeuille des entreprises d'assurances sont principalement composés d'obligations, la modélisation du crédit est devenue indispensable.

### 6.3.1 Analyse descriptive

Nous choisissons dans ce modèle l'indice ITRAXX représentant le spread de crédit des dérivés de crédit CDS <sup>18</sup> sur des obligations. Mais ce choix doit être effectué en adéquation avec la **duration** moyenne du portefeuille de l'entreprise d'assurance.

Ci-dessous la série des spreads de l'indice ITRAXX <sup>192</sup> de maturité 5 ans des obligations privées de nature "Investment Grade".

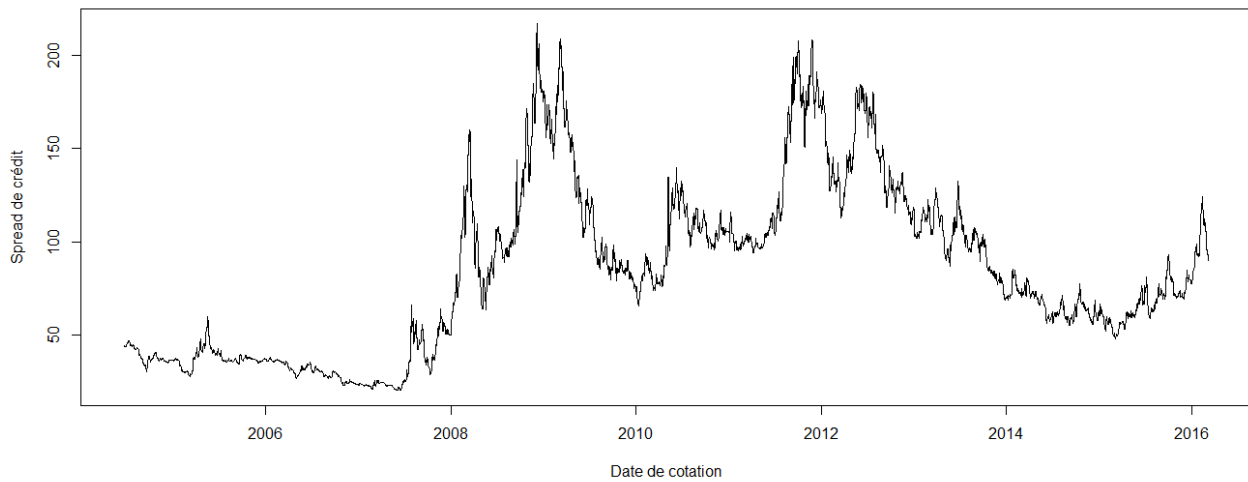


FIGURE 6.17 – Évolution du spread de crédit ITRAXX en bps sur la période 22-06-2004 → 28-02-2016

---

18. Credit Default Swap.

19. Source Bloomberg - TICKER "ITRXXE CBIL Curncy"

### 6.3.2 Modélisation CIR

Pour modéliser le défaut, on utilise souvent le modèle de crédit à **intensité**; la loi du temps de défaut est donnée par :

$$\mathbb{P}\{\tau < t\} = 1 - e^{-\lambda t}$$

où  $\lambda$  est appelé **intensité**.

Dans un cadre *simple*, on valorise sous la probabilité risque-neutre la **jambe fixe** et la **jambe variable**, assujetties au risque de défaut, d'un CDS avec un **taux de recouvrement** (constant)  $R$  :

$$\begin{aligned} JF(s) &:= s \sum_{i=0}^{n-1} (t_{i+1} - t_i) \mathbb{E}[e^{-rT_{i+1}} 1_{\{\tau > T_{i+1}\}}] \\ &= s \sum_{i=0}^{n-1} (t_{i+1} - t_i) e^{-(r+\lambda)T_{i+1}} \\ &\simeq s \int_0^t e^{-(r+\lambda)u} du \\ &= s \frac{1 - e^{-(r+\lambda)T}}{r + \lambda} \end{aligned}$$

$$\begin{aligned} JV &:= (1 - R) \mathbb{E}[e^{-r\tau} 1_{\{\tau \leq T\}}] \\ &= (1 - R) \int_0^T \lambda e^{-(r+\lambda)u} du \\ &= (1 - R) \frac{\lambda(1 - e^{-(r+\lambda)T})}{r + \lambda} \end{aligned}$$

La valeur du spread (*fair*) est déterminée tel que  $JF(s) = JV$  i.e.

$$s = (1 - R)\lambda \tag{6.11}$$

Dans la littérature, on choisit souvent un processus CIR<sup>20</sup> pour l'intensité  $\lambda$ . Du fait de la relation (6.11) nous **choisissons** donc également un processus **CIR** pour le **spread** de crédit :

$$ds_t = \kappa_s(\mu_s - s_t)dt + \sqrt{s_t}\sigma_s dW_t \tag{6.12}$$

avec

- $\mu_s$  le retour à la moyenne
- $\kappa_s$  la force de rappel
- $\sigma_s$  la volatilité

---

20. Cox-Ingersoll-Ross

La **discrétisation** de ce processus continu avec un schéma de **Milstein**<sup>21</sup> d'un pas  $\delta = 1$  est donné par

$$s_{t+1} - s_t = \left( \kappa_s(\mu_s - s_t) - \frac{\sigma_s^2}{4} \right) + \sigma_s \sqrt{s_t} \epsilon_t + \frac{\sigma_s^2}{4} \epsilon_t^2 \quad (6.13)$$

Si on applique le lemme d'Itô à la diffusion (6.12) avec  $\tilde{s}_t = \sqrt{s_t}$  on a l'EDS suivante<sup>22</sup> :

$$d\tilde{s}_t = \frac{1}{2\tilde{s}_t} \left( \kappa_s(\mu_s - \tilde{s}_t^2) - \frac{\sigma_s^2}{4} \right) dt + \frac{\sigma_s}{2} dW_t \quad (6.14)$$

Si l'on effectue maintenant une discrétisation selon le schéma d'**Euler**, on a :

$$\Delta\tilde{s}_t := \tilde{s}_{t+1} - \tilde{s}_t = \frac{1}{2\tilde{s}_t} \left( \kappa_s(\mu_s - \tilde{s}_t^2) - \frac{\sigma_s^2}{4} \right) + \frac{\sigma_s}{2} \epsilon_t \quad (6.15)$$

Remarquons avec  $f(x) := x^2$

$$\begin{aligned} f(\tilde{s}_t + \Delta\tilde{s}_t) - f(\tilde{s}_t) &= (\tilde{s}_t + \Delta\tilde{s}_t)^2 - \tilde{s}_t^2 \\ &= (\Delta\tilde{s}_t)^2 + 2\tilde{s}_t\Delta\tilde{s}_t \\ &\simeq \frac{\sigma_s^2}{4} \epsilon_t^2 + \left( \kappa_s(\mu_s - \tilde{s}_t^2) - \frac{\sigma_s^2}{4} \right) + \tilde{s}_t\sigma_s\epsilon_t \end{aligned}$$

On retrouve le schéma de **Milstein** (6.13) avec  $\tilde{s}_t = \sqrt{s_t}$ .

Ainsi, à partir de l'équation (6.15), on trouve donc finalement les paramètres  $\theta := (\kappa_s, \mu_s, \sigma_s)$  du processus CIR, en effectuant une **régression** des moindres carrés ordinaires :

$$y_t = \alpha_1 x_t^1 + \alpha_2 x_t^2 + \tilde{\epsilon}_t \quad (6.16)$$

avec

$$\begin{cases} y_t := \tilde{s}_{t+1} - \tilde{s}_t \\ x_t^1 := \tilde{s}_t^{-1} \\ x_t^2 := \tilde{s}_t \\ \alpha_1 := \frac{4\kappa_s\mu_s - \sigma_s^2}{8} \\ \alpha_2 := -\frac{\kappa_s}{2} \end{cases}$$

et  $\{\tilde{\epsilon}_t, t \in \mathbb{Z}\}$  est un bruit blanc gaussien de variance  $\tilde{\sigma}_s := \frac{\sigma_s}{2}$ .

Ainsi on obtient les estimateurs des paramètres du modèle en résolvant un problème de moindres carrés ordinaires :

$$\begin{aligned} \hat{\theta} &= (\kappa_s, \mu_s, \sigma_s) \\ &\equiv \left( \underset{\alpha_1, \alpha_2}{\operatorname{argmin}} \sum_{i=2}^H (y_{t-i} - \alpha_1 x_{t-i}^1 - \alpha_2 x_{t-i}^2)^2, \hat{\sigma}_s \right) \\ &= \left( -2\hat{\alpha}_2, \frac{4\hat{\alpha}_1 + \hat{\sigma}_s}{-4\hat{\alpha}_2}, 2\hat{\sigma}_s \right) \end{aligned}$$

21. Cf. l'article de F. Planchet & P-E Thérond.

22. Cf. l'article d'A. Alfonsi [4] page 3.



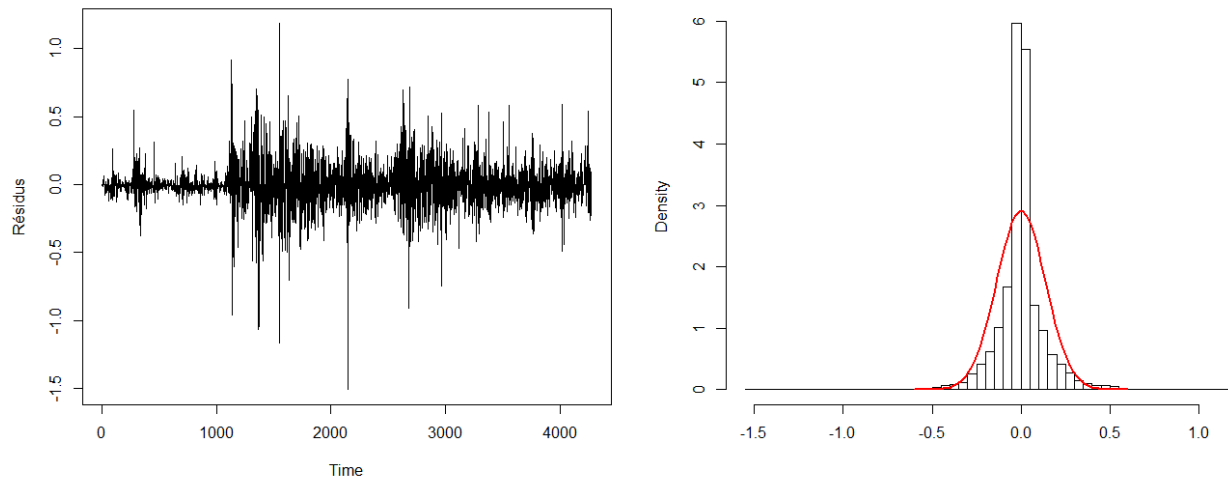


FIGURE 6.18 – Distribution empirique des résidus du modèle CIR sur la période 22-06-2004  $\leftrightarrow$  28-02-2016

## 6.4 Modèle dynamique pour l'inflation

L'inflation constitue un **risque d'actif** directement si l'entreprise d'assurance détient des obligations indexées sur l'inflation ou indirectement du fait des interactions macro-économiques avec les autres type d'actifs (taux d'intérêt, action, ...).

L'inflation constitue également un **risque de passif**. En effet, pour les risques à développement **long** on les exemples suivants :

- En assurance Vie : revalorisation des rentes en prévoyance ou conversion en rente des contrats de capitalisation, ...
- En assurance non-Vie : valeur de l'exposition sous risque qui suit peu ou prou l'inflation, revalorisation des rentes en responsabilité civile, ...

La couverture de l'inflation est un problème complexe en effet le marché de l'inflation est **asymétrique** : il y a plus d'acheteurs que de vendeurs d'inflation

### 6.4.1 Analyse descriptive

La **mesure** de l'inflation au mois  $m$  sur une période de  $k$  mois est la variation définie à partir du log-rendement de l'Indice<sup>23</sup> des Prix à la Consommation (IPC<sup>24</sup>) :

$$i_{m,k} := \ln \frac{\text{IPC}_m}{\text{IPC}_{m-k}}$$

Si  $k=1$ , la variable d'observation est le taux d'inflation mensuel que l'on peut annualiser :

$$i_m^a := (1 + i_{m,1})^{12} - 1$$

Néanmoins, pour éviter la problématique de **saisonnalité**<sup>25</sup> de l'inflation, la variable d'observation est souvent l'**inflation annuelle**  $i_{m,12}$  que l'on calcule chaque mois<sup>26</sup> après publication de l'indice des prix. Nous omettrons par la suite l'indice 12.

---

23. Il mesure l'évolution de la valeur d'un panier de biens représentatif de la consommation. Il est publié mensuellement par l'INSEE.

24. CPI en anglais.

25. Demande d'énergie en hiver, Soldes, ...

26. Inflation glissante.

Ci-dessous, le graphique représente l'évolution mensuelle du taux annuel d'inflation française  $i_m$ . Nous pouvons observer en particulier une rupture de la dynamique à partir des années 2009-2010.

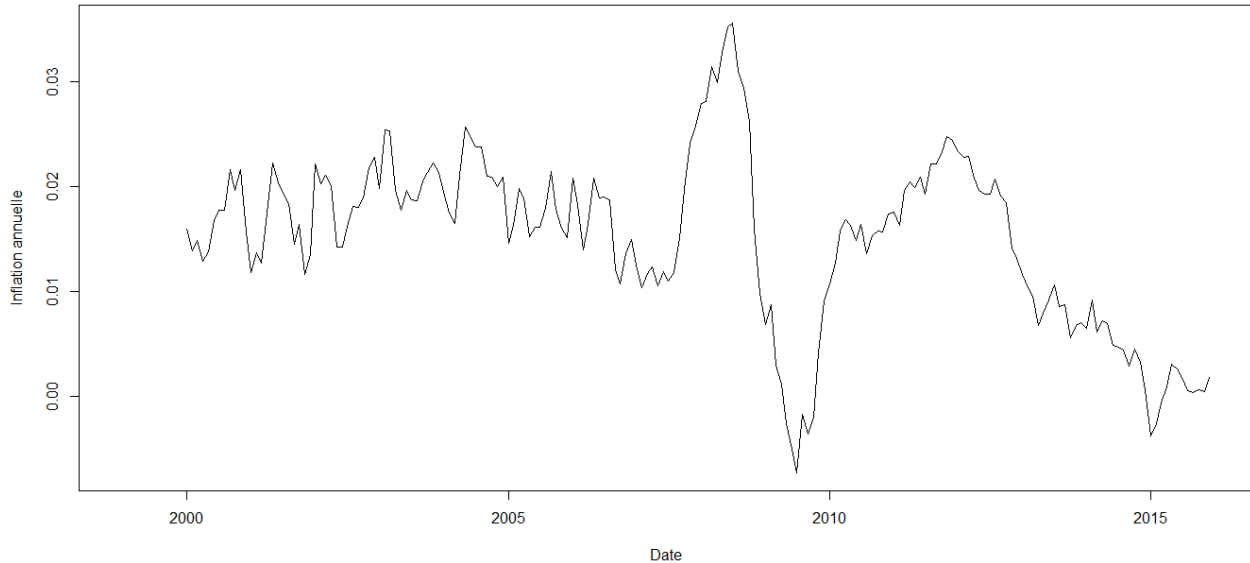


FIGURE 6.19 – Evolution du taux d'inflation annuelle glissante en France sur la période janv. 1999  $\mapsto$  dec. 2015

## 6.4.2 Modélisation auto-régressive

En monde **risque-neutre** le taux instantané<sup>27</sup> d'inflation, par analogie aux taux d'intérêt, est souvent modélisée avec une dynamique de type **Vacisek** c'est-à-dire avec un processus d'**Orstein-Uhlenbeck** (OU) :

$$di_t = \kappa(\mu - i_t) dt + \sigma dW_t^{\mathbb{Q}}$$

Or les processus OU sont **stables** par changement de probabilité. Il nous semble donc **naturel** de choisir un processus OU pour la dynamique du taux d'inflation en **monde réel**.

L'utilisation d'un processus de **retour** à la moyenne est un critère qui semble intéressant dans le cadre de projections à **long** terme. Le paramètre de moyenne à long terme ( $\mu$ ) devient interprétable. Il ne semble pas viable économiquement une inflation explosive. Rappelons également que les processus OU permettent des trajectoires dans les territoires négatifs<sup>28</sup>.

Pour des simulations monde réel, on retrouve ainsi ce type de processus, par exemple, dans le GSE d'**Ahlgrim**<sup>29</sup>

27. i.e. dans la définition du log-rendement le  $lag\ k \rightarrow 0$

28. Déflation.

29. Cf. l'article [3]. Ce GSE est présenté dans le cadre du chapitre sur les simulations multi-dimensionnelles page 80.

ou dans celui de Brennan et Xia<sup>30</sup> pour modéliser le taux d'inflation instantané à l'instant  $t$ .

Pour **simuler**, il est nécessaire de **discrétiser**<sup>31</sup> le processus continu. Le processus OU discrétisé est un processus *auto-régressif* de type AR(1) de paramètres  $\alpha_1, \alpha_2, \sigma_\epsilon$  :

$$i_{m+1} = i_m e^{-\kappa} + \mu(1 - e^{-\kappa}) + \sqrt{\frac{1 - e^{-2\kappa}}{2\kappa}} \sigma \epsilon_m \quad (6.17)$$

$$:= \alpha_1 + \alpha_2 i_m + \tilde{\sigma} \epsilon_m \quad (6.18)$$

où  $\tilde{\epsilon}$  est un bruit blanc gaussien.

Nous avons par identification :

$$\begin{aligned} \alpha_2 &:= e^{-\kappa} \\ \alpha_1 &:= \mu - (1 - e^{-\kappa}) \\ \tilde{\sigma} &:= \sigma \sqrt{\frac{1 - e^{-2\kappa}}{2\kappa}} \end{aligned}$$

c'est-à-dire les paramètres du processus OU (6.17) sont estimés via l'estimation standard du processus AR(1) (6.18) :

$$\begin{aligned} \hat{\kappa} &= -\ln \hat{\alpha}_2 \\ \hat{\mu} &= \frac{\hat{\alpha}_1}{1 - \hat{\alpha}_2} \\ \hat{\sigma} &= \frac{\tilde{\sigma}}{\sqrt{\frac{1 - e^{\hat{\kappa}}}{2\hat{\kappa}}}} \end{aligned}$$

On retrouve ce type de modélisation auto-régressive sans faire forcément explicitement référence à une discrétisation d'un processus OU par exemple dans le GSE issu du modèle de Wilkie<sup>32</sup> où le taux d'inflation<sup>33</sup> anglais est modélisé selon un processus<sup>34</sup> AR(1) :

$$i_n = \bar{\iota} + \alpha(i_{n-1} - \bar{\iota}) + \sigma \epsilon_n$$

où  $\iota$  représente l'inflation long-terme.

Néanmoins, plusieurs critiques ont été apportées à utiliser un modèle **standard** auto-régressif. Par exemple Kitts et Geoghegan dans leur article<sup>35</sup> respectifs indiquent que les tests statistiques dans le cadre du modèle de Wilkie montrent une dépendance et une non-normalité des résidus, suggérant une **non-linéarité**.

---

30. Cf. l'article []

31. Discrétisation d'Euler.

32. Cf. l'article [39]. Et également Cf. le paragraphe sur le modèle de Wilkie dans le cadre du chapitre sur les simulations multidimensionnelles page 80.

33. log-rendement annuel du Retail Price Index (RPI).

34. Dans sa version de 1986.

35. Cf. les articles [20] et [24].

Si nous effectuons un test de stationnarité de **Dickey-Fuller** sur la série des données françaises (Cf. graphique 6.19 page 131),

- sur la période 2000  $\mapsto$  2009, la p-value est inférieure à 10% on pourrait rejeter l'hypothèse nulle de non-stationnarité.
- sur la période 2010  $\mapsto$  2015, la p-value est très supérieure au seuil critique on ne peut donc rejeter l'hypothèse de non- stationnarité.

Ainsi cela nous suggère un modèle non-linéaire. C'est l'objet de la section suivante.

### 6.4.3 Modélisation SETAR

La Banque Centrale Européenne a comme **objectif**<sup>36</sup> une inflation annuelle (**Harmonised Index Of Consumer Prices**) inférieure à **2%** mais proche de 2% à moyen-terme.

La banque centrale européenne adaptera donc sa **politique** monétaire **en fonction** de l'environnement d'**inflation**. Sa politique sera différente si l'inflation est basse (**moins** de 2%) ou si l'inflation est forte (**plus** de 2%). Ainsi la dynamique de l'inflation sera **différente** en fonction de son niveau.

Dans une modèle OU, on peut par exemple comparer l'estimation du paramètre de moyenne à long terme  $\mu$  avec cet objectif.

Néanmoins, la dynamique dans un OU n'est pas dépendant de son spread par rapport à l'objectif de la BCE. Or il semble par exemple que les interventions d'assouplissement quantitatif<sup>37</sup> montrent un comportement différencié de la Banque centrale pour réguler l'inflation en fonction de son niveau.

L'idée, ici, est donc d'utiliser un modèle de type SETAR pour la variable  $i_m$  avec comme paramètre de seuil  $c = 2\%$  :

$$i_m = (\phi_0 + \sum_{i=1}^p \phi_i i_{m-i}) 1_{\{q_{m-1} \leq c\}} + (\phi'_0 + \sum_{i=1}^p \phi'_i i_{m-i}) 1_{\{q_{m-1} > c\}} + \epsilon_m \quad (6.19)$$

Dans la fonction indicatrice, nous choisissons un *lag*<sup>38</sup> de 1 de la variable inflation  $q_{t-1} := i_{m-1}$ , nous supposons ainsi ici une forte réactivité de la Banque centrale. Mais idéalement, il faudrait prendre un *lag* de réaction. On pourrait remplacer la variable seuil de la fonction indicatrice  $1_{\{q_{m-1} \leq c\}}$  par une moyenne mobile de l'inflation.

Considérons l'historique de fréquence mensuelle  $\{i_{m-H+1}, \dots, i_{m-1}\}$  de l'indice des prix à la consommation sur la période janv.2000  $\mapsto$  dec. 2015.

---

36. Cf. The ECB's monetary policy strategy.

37. Quantitative easing.

38. Le paramètre  $d$  de la fonction indicatrice  $1_{\{i_{m-d} \leq c\}}$  est aussi appelé *delay parameter*.

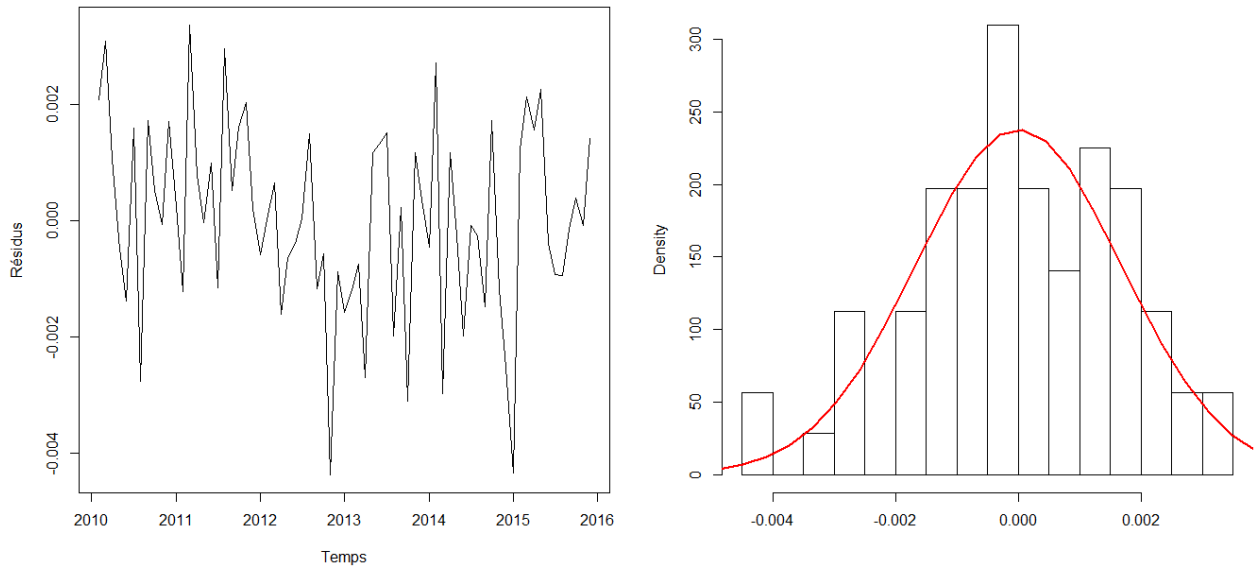


FIGURE 6.20 – Distribution empirique des résidus du modèle SETAR sur la période janv. 2000  $\mapsto$  déc. 2015

## 6.5 Modèle dynamique pour l'immobilier

La mesure **mensuelle** du rendement de l'immobilier est définie comme pour l'inflation à partir des log-rendements annuels<sup>39</sup> de l'Indice des Prix des Logements par :

$$re_m := \ln \frac{IPL_m}{IPL_{m-12}}$$

### 6.5.1 Analyse descriptive

L'INSEE fournit un historique **trimestrielle** d'indice des prix des logements dans l'ancien :

- France métropolitaine sur la période trimestre 1 1999  $\Rightarrow$  trimestre 3 2015.
- Paris sur la période trimestre 1 1992  $\Rightarrow$  trimestre 3 2015.

Nous n'avons pas donc pas un historique à fréquence **mensuelle**, on calcule donc le log-rendement annuel<sup>40</sup> à fréquence **trimestrielle** par la transformation :

$$re_{tr} := \ln \frac{IPL_{tr}}{IPL_{tr-4}}$$

Nous remarquons en particulier sur la période récente 2009-2012 de **forts** mouvements de l'indice immobilier et de son log-rendement annuel sur les deux périmètres respectifs.

---

39. Un lag de 12 mois pour la problématique de saisonnalité.

40. Du fait de la saisonnalité.

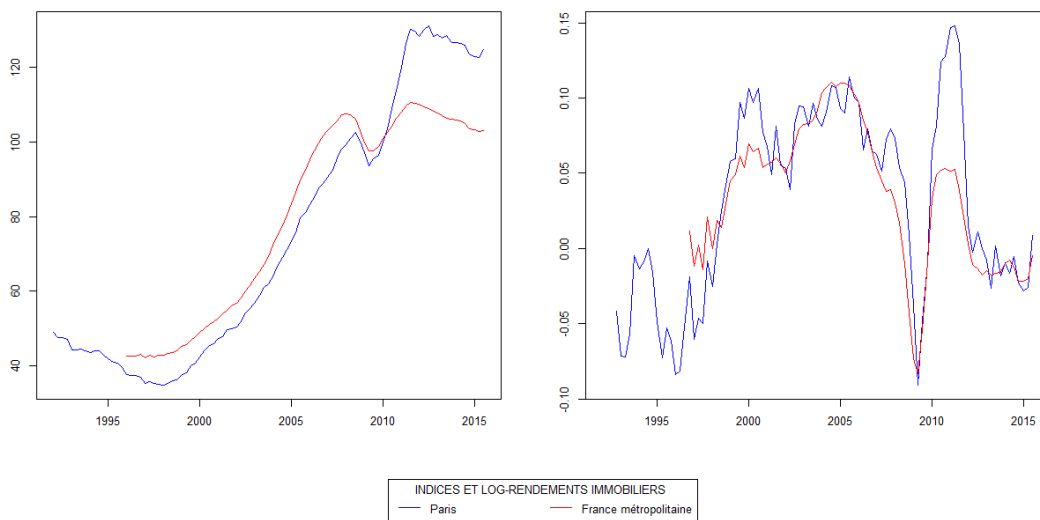


FIGURE 6.21 – Evolution sur la période trimestre 1 1992  $\rightarrow$  trimestre 3 2015



## 6.5.2 Modélisation BS

Pour des simulations risque-neutre, le modèle Black-Sholes i.e. un brownien géométrique est choisi pour modéliser la dynamique des rendements immobilier comme pour les actions.

Ce n'est pas cette modélisation que nous privilégions.

## 6.5.3 Modélisation OU

Une modélisation que l'on rencontre souvent dans la littérature<sup>41</sup>, en particulier pour le monde réel, est de type Ornstein-Uhlenbeck.

On peut donc modéliser le log-rendement immobilier<sup>42</sup> par un processus **auto-régressif** d'ordre **1**, comme pour l'inflation<sup>43</sup> dans le cadre d'une modélisation auto-régressive (**OU discrète**<sup>44</sup>) :

$$re_m = re_{m-1}e^{-\kappa_{re}} + \mu_{re}(1 - e^{-\kappa_{re}}) + \sqrt{\frac{1 - e^{-2\kappa_{re}}}{2\kappa_{re}}} \sigma_{re} \epsilon_m^{re}$$

où  $\epsilon_m^{re}$  est un bruit blanc de variance **unitaire**.

On trouve les coefficients  $\kappa_{re}, \mu_{re}, \sigma_{re}$  par régression<sup>4</sup> linéaire simple de type AR(1).

En prenant en compte la période récente de **crises**, la fonction d'auto-corrélation partielle montre qu'un modélisation AR d'ordre **1** n'est **plus** adéquate, les auto-corrélations partielles d'ordre supérieur à 1 ne sont plus significativement non nulle.

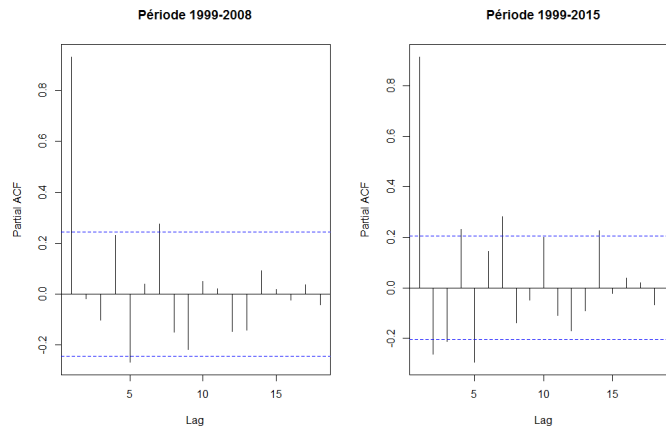


FIGURE 6.22 – Auto-corrélations partielles en fonction de la période considérée.

---

41. Modèle d'Ahlgrim par exemple.

42. re comme real estate.

43. Cf. section 6.4.2 page 131.

44. re indicé par m avec un pas de discrétisation de 1 représentant le mois.

# Conclusion

Le générateur de scénarios économiques que nous avons construit est un GSE sous probabilité **historique** multi-dimensionnel de type composite dont la structure de corrélation est définie

- **historiquement** via un tirage **bootstrap vectoriel**<sup>45</sup> de l'ensemble des actifs considérés.
- via la copule de **Paul Deheuvels** pour la structure de **dépendance** des **queues** des distributions **marginales** des différents actifs.

Cette méthode à l'avantage d'être **pérenne** quel que soit le choix économétrique de chaque actif et le **nombre** d'actifs modélisés et ne nécessite pas d'estimation de copules qui est souvent peu aisée en grande dimension.

La **cohérence** économique est déterminée via le choix économétrique pour chaque variable d'intérêt et principalement déterminée via un tirage *bootstrap vectoriel*.

Retenons également pour les taux d'intérêt le choix d'introduire un modèle économétrique directement au niveau des paramètres de la fonctionnelle de **Nelson-Siegel**. Ce choix nous semble judicieux du fait que par une nature "**structure par terme**" les taux d'intérêt sont une variable multidimensionnelle. Il conserve les propriétés de dépendance entre les **différentes maturités** et donc le caractère d'absence d'opportunité d'**arbitrage**.

Les étapes qui nécessitent une attention **particulière** sont le **choix** du modèle **économétriques**. Ce choix peut évoluer avec le temps. En effet, on observe notamment avec les **crises** de ces dernières années des changements de comportement des variables économiques et financières.

---

45. Donc à une même date.

# Annexe A

## Copules

### A.1 Définition

On appelle **copule bi-dimensionnelle** la **fonction**  $C$  à deux variables qui vérifie les propriétés suivantes :

1.  $\text{Dom } C = [0, 1] \times [0, 1]$  i.e. le support de la copule est le carré unité.
2.  $\forall u \in [0, 1] C(u, 0) = C(0, u) = 0$  i.e. la fonction est nulle dès que l'une des variables aléatoires est nulle<sup>1</sup>.
3.  $\forall u \in [0, 1] C(u, 1) = C(1, u) = u$  i.e. les marges sont **uniformes**.
4.  $\forall (u_1, v_1, u_2, v_2) \in [0, 1]^2 \times [0, 1]^2$  tels que  $0 \leq u_1 \leq v_1$ , tels que  $0 \leq u_2 \leq v_2$   
 $C(v_1, v_2) - C(v_1, u_2) - C(v_2, u_1) + C(u_1, u_2) \geq 0$  i.e. la **probabilité**  $\mathbb{P}\{u_1 \leq U_1 \leq v_1, u_2 \leq U_2 \leq v_2\}$  est positive<sup>2</sup>.

Une copule bi-dimensionnelle est donc une **distribution** de probabilité 2D avec des marginales  $(U_1, U_2)$  **uniformes**, :

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

Nous parlons d'expression plus *simple* car il est possible de **séparer** le comportement **marginal** des variables aléatoires, de la structure de **dépendance** proprement dite. C'est l'objet du théorème fondamental de Sklar :

#### **Théorème A.1.1 (Théorème de Sklar)**

Si  $C$  est une **copule** et  $F_1, \dots, F_d$  les fonctions de répartition **univariées** de  $d$  variables aléatoires  $X_i$  alors la loi **jointe**  $F$  peut s'écrire :

$$\forall x \in \mathbb{R}^d \quad F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad (\text{A.1})$$

---

1. En anglais on parle de *grounded function* (clouée au sol).  
2. En anglais on parle de la propriété *2-increasing*.

**Réciproquement**, si  $F$  est une distribution **jointe** et si les distributions marginales  $F_i$  sont **continues**, il existe une copule **unique** satisfaisant l'équation (A.1) :

$$\forall u \in [0, 1]^d \quad C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)) \quad (\text{A.2})$$

La première partie du théorème nous indique que l'on peut **construire** une distribution multivariée à partir de marginales et d'un choix de copule.

La deuxième partie du théorème nous indique que le problème de dépendance multivariée peut se décomposer en deux sous-problématiques :

- **identifier** le comportement marginal de chaque variable aléatoire.
- **analyser** de la structure de la dépendance.

Une fois la copule déterminée, la simulation de  $d$  variables aléatoires **liées** de distributions respectives  $F_1, \dots, F_d$  s'effectue en deux étapes :

1. **simuler** une réalisation d'un vecteur  $U$  de  $\mathbb{R}^d$  de distribution la copule  $C$ .
2. **obtenir** la réalisation du vecteur  $X$  en utilisant la transformation <sup>3</sup> :

$$X := (x_1, \dots, x_n)^T := \left( F_1^{-1}(u_1), \dots, F_d^{-1}(u_d) \right)$$

La **difficulté** est donc dans la simulation de la copule. De manière générale, on peut écrire en dimension deux <sup>4</sup> :

$$\begin{aligned} C(u_1, u_2) &= \mathbb{P}\{U_1 \leq u_1, U_2 \leq u_2\} \\ &= \mathbb{P}\{U_2 \leq u_2 \mid U_1 \leq u_1\} \mathbb{P}\{U_1 \leq u_1\} \end{aligned} \quad (\text{A.3})$$

d'après le théorème de Bayes

Ainsi, la décomposition (A.3) nous indique que la simulation de deux variables  $U_1$  et  $U_2$  selon la loi jointe définie par la copule  $C$ , peut s'effectuer en simulant <sup>5</sup> :

- une variable uniforme  $U_1 : u_1 = u_1^\perp$
- simulant une variable  $U_2 \mid U_1 = u_1$  selon la loi **conditionnelle** :

---

3. En effet si  $X$  est une variable aléatoire de distribution  $F$  alors  $F(X)$  est une variable aléatoire **uniforme**.

4. La généralisation en dimension  $d$  s'effectue de la même façon.

5.  $U_1^\perp$  et  $U_2^\perp$  sont deux variables aléatoires **uniformes indépendantes** de réalisation respectives  $u_i^\perp$

$$\begin{aligned}
C_{2|1}(u_2; u_1) &:= \mathbb{P}\{U_2 \leq u_2 \mid U_1 = u_1\} \\
&= \lim_{\delta u \rightarrow 0^+} \frac{\mathbb{P}\{U_2 \leq u_2, u_1 \leq U_1 \leq u_1 + \delta u\}}{\mathbb{P}\{u_1 \leq U_1 \leq u_1 + \delta u\}} \\
&= \lim_{\delta u \rightarrow 0^+} \frac{C(u_1 + \delta u, u_2) - C(u_1, u_2)}{\delta u} \\
&\quad \text{d'après les propriétés de Sklar de la fonction copule (A.1) page 139.} \\
&= \frac{\partial C}{\partial u_1}(u_1, u_2)
\end{aligned}$$

c'est-à-dire une réalisation  $U_2 \mid U_1 = u_1^\perp$  est donnée<sup>6</sup> par  $u_2 = C_{2|1}(\bullet; u_1^\perp)^{-1}(u_2^\perp)$ .

Pour terminer, la simulation de deux variables  $X_1$  et  $X_2$  de loi jointe définie par la copule  $C$ , consiste à effectuer les transformations réciproques :

$$X_1 := F_1^{-1}(u_1) \quad X_2 := F_2^{-1}(u_2)$$

Néanmoins, cette démarche est simple<sup>7</sup> si l'on peut déterminer **analytiquement** la **réci-proque** de la distribution **conditionnelle**.

## A.2 Copule paramétrique elliptique

Avant de présenter succinctement les distributions elliptiques, revenons sur la simulation d'un vecteur gaussien **corrélé** dont la structure de dépendance est définie par la matrice de **corrélations**.

Dans le cas bivarié, si  $X = (X_1, X_2)^T$  est un vecteur de  $\mathbb{R}^2$  de moyenne  $\mu$  et de matrice **variance-covariance**<sup>8</sup>  $\Sigma$  alors  $X$  a la même loi que le vecteur  $\mu + PZ$  où

- $Z$  est un vecteur aléatoire de  $\mathbb{R}^2$  à composantes **indépendantes** et **gaussiennes** centrées réduites.
- $P$  est la matrice<sup>9</sup> pseudo-racine carrée de  $\Sigma$  issue de la **décomposition de Cholesky** :

$$\Sigma := \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \Rightarrow \Sigma = P P^T \Rightarrow \text{avec } P := \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1 - \rho^2} \end{pmatrix}$$

Ainsi pour **simuler** ce vecteur  $X$  avec des composantes **corrélées**, il suffit de simuler deux variables aléatoires  $Z_1$  et  $Z_2$  gaussiennes indépendantes et d'utiliser la décomposition de **Cholesky** :

$$\begin{aligned}
X_1 &= \mu_1 + Z_1 \\
X_2 &= \mu_2 + \rho Z_1 + \sqrt{1 - \rho^2} Z_2
\end{aligned}$$

6. Cf. la note (de bas de page) 1 précédente.

7. Pour les copules paramétriques, il existe des méthodes alternatives de simulation en fonction de la famille de copule, en particulier pour les copules archimédiennes.

8. Matrice semi-définie positive.

9. Matrice triangulaire inférieure.

Cet algorithme est très utilisé pour simuler, par exemple, des mouvements **browniens corrélés**. Nous allons retrouver cet algorithme avec la simulation de la copule gaussienne selon la méthode de distributions conditionnelles décrite précédemment. C'est la copule paramétrique la plus populaire, elle est définie par

$$\forall u \in [0, 1]^d \quad C(u_1, \dots, u_d) := \Phi_R(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))$$

où

- $\Phi^{-1}(\cdot)$  la réciproque de la distribution normale centrée réduite **univariée**.
- $\Phi_R(\cdot)$  la fonction de distribution de la loi normale centrée réduite multivariée et  $R$  sa matrice de **corrélation**<sup>10</sup> :

En dimension 2 :

$$C(u_1, u_2; \rho) := \int_0^{\Phi^{-1}(u_1)} \int_0^{\Phi^{-2}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{x^2 + y^2 - 2\rho xy}{2(1-\rho^2)}\right\} dx dy$$

On peut écrire cette copule sous la forme suivante de J-F. Jouanin en utilisant seulement la distribution normale centrée réduite **univariée** :

$$C(u_1, v_1; \rho) = \int_0^{u_1} \Phi\left(\frac{\Phi^{-1}(u_2) - \rho \Phi^{-1}(x)}{\sqrt{1-\rho^2}}\right) dx$$

Ainsi le calcul de la distribution **conditionnelle** se calcule facilement :

$$\begin{aligned} C_{2|1}(u_2; u_1, \rho) &= \frac{\partial C}{\partial u_1}(u_1, u_2) \\ &= \Phi\left(\frac{\Phi^{-1}(u_2) - \rho \Phi^{-1}(u_1)}{\sqrt{1-\rho^2}}\right) \end{aligned} \tag{A.4}$$

La simulation d'une copule gaussienne consiste donc à simuler en utilisant la démarche de loi conditionnelle itérée décrite 140 et la forme de (A.4) :

- une loi gaussienne centrée réduite  $x_1 = \Phi^{-1}(u_1) = \Phi^{-1}(u_1^\perp)$
- une loi gaussienne centrée réduite telle que :

$$\begin{aligned} x_2 &= \Phi^{-1}(u_2) = \sqrt{1-\rho^2} \Phi^{-1}(u_2^\perp) + \rho \Phi^{-1}(u_1^\perp) \\ &= \rho x_1 + \sqrt{1-\rho^2} \Phi^{-1}(u_2^\perp) \end{aligned}$$

On **retrouve** la méthode de simulation par la factorisation de **Cholesky**.

Utiliser une **matrice** de corrélation pour lier des variables gaussiennes consiste donc **implicitement** à supposer une structure de dépendance définie par une copule **gaussienne** avec des marginales gaussiennes.

---

10. Dans le cas centré réduit la matrice de variance-covariance est la matrice de corrélation.

Revenons sur un point souvent passé sous silence dans l'approche de simulation par distribution conditionnelle.

$$\Phi_2(X_1, X_2; \rho) = \Phi(X_1) \mathbb{P}\{X_2 \leq x_2 \mid X_1 \leq x_1\}$$

Pour le deuxième terme on considère on considère plutôt la loi conditionnelle

$$\mathbb{P}\{X_2 \leq x_2 \mid X_1 = x_1\}$$

que la loi

$$\mathbb{P}\{X_2 \leq x_2 \mid X_1 \leq x_1\}$$

Dans le cas de la loi gaussienne bivariee la loi de  $X_2 \mid X_1 = x_1$  est normale alors que  $X_2 \mid X_1 \leq x_1$  ne l'est pas. Néanmoins, cette dernière peut être approximée<sup>11</sup> par une loi normale avec la même moyenne et variance.

Du fait de la **simplicité** de la simulation et du fait que le paramètre de la copule est **interprétable** par la mesure de corrélation usuelle, la copule gaussienne est très utilisées dans les problématiques de dépendances multivariées en finance et en assurance.

On **observe** souvent dans les données financière une corrélation des extrêmes. On **souhaite** donc que la structure de dépendance définie par la copule présente des dépendances de queues.

De façon générale, on peut exprimer les coefficients de dépendance de queues en fonction de la copule :

$$\begin{aligned} p^{\max} &:= \lim_{\alpha \rightarrow 1^-} \mathbb{P}\{X_1 > F_1^{-1}(\alpha) \mid X_2 > F_2^{-1}(\alpha)\} \\ &= \lim_{u \rightarrow 1^-} \mathbb{P}\{U_1 > u \mid U_2 > u\} \\ &= \lim_{u \rightarrow 1^-} \frac{1 - \mathbb{P}\{U_1 \leq u\} - \mathbb{P}\{U_2 \leq u\} + \mathbb{P}\{U_1 \leq u, U_2 \leq u\}}{1 - \mathbb{P}\{U_2 \leq u\}} \\ &= \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u} \end{aligned}$$

et de la même façon on trouve

$$\begin{aligned} p^{\min} &= \lim_{\alpha \rightarrow 0^+} \mathbb{P}\{X_1 \leq F_1^{-1}(\alpha) \mid X_2 \leq F_2^{-1}(\alpha)\} \\ &= \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u} \end{aligned}$$

Cependant, en pratique, en particulier pour la copule gaussienne ces coefficients **ne sont pas facilement calculable**. Néanmoins on peut remarquer :

$$\begin{aligned} \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u} &= \lim_{u \rightarrow 0^+} \frac{C(u, u) - C(0, 0)}{u - 0} \\ &= \lim_{\alpha \rightarrow 0^+} \frac{dC(u, u)}{du} \end{aligned}$$

---

11. Pour une discussion plus détaillée Cf. l'article [27].

et la différentielle totale s'écrit :

$$dC(u, u) = \partial_1 C(u, u) du + \partial_2 C(u, u) du$$

On a vu précédemment que  $\partial_i C(u_1, u_2)$  représente la loi conditionnelle et en utilisant la symétrie de la copule gaussienne on a :

$$p_\alpha^{min} = 2 \times \lim_{u \rightarrow 0^+} \mathbb{P}\{U_2 \leq u \mid U_1 = u\}$$

et en utilisant de nouveau la forme (A.4) page 142 et utilisant les propriétés de la fonction de répartition :

$$\begin{aligned} p_\alpha^{min} &= 2 \times \lim_{u \rightarrow 0^+} \Phi\left(\frac{\Phi^{-1}(u)(1-\rho)}{\sqrt{1-\rho^2}}\right) \\ &= 0 \end{aligned}$$

De la même façon, on trouve également  $p_\alpha^{max} = 0$ . La copule gaussienne ne présente donc **pas de dépendances** des extrêmes

La copule de **Student** de degrés de liberté  $\nu$  et de matrice de corrélation  $R$  définie par

$$\forall u \in [0, 1]^d \quad C(u_1, \dots, u_d) := t_{R, \nu}\left(t_\nu^{-1}(u_1), \dots, t_\nu^{-1}(u_d)\right)$$

est une distribution elliptique qui permet d'**introduire** des **dépendances** de queues<sup>12</sup> **symétriques** entre les variables :

$$p_\alpha^{min, max} := 2 - 2t_{\nu+1}\left(\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right) > 0$$

### A.3 Copule paramétrique archimédienne

Quant aux copules **archimédiennes**, elles permettent d'introduire des dépenses de queues **asymétriques** c'est-à-dire les dépendances de queues à **gauche** et à **droite différentes**.

Une copule **archimédienne**<sup>13</sup> est une copule construite à partir d'une fonction  $\varphi_\theta$  appelé **générateur** paramétrique :

$$\forall u \in [0, 1]^d \quad C_\theta(u) := \varphi_\theta^{-1}\left(\varphi_\theta(u_1) + \dots + \varphi_\theta(u_d)\right)$$

vérifiant les propriétés suivantes :

- $\sum_{i=1}^d \varphi_\theta(u_i) \leq \varphi_\theta(0)$
- $\varphi_\theta$  est de classe  $C^2$  telle que  $\varphi_\theta(1) = 0$ ,  $\varphi_\theta^{(1)} \leq 0$  et  $\varphi_\theta^{(1)} \geq 0$

On peut citer comme exemple, les copules de :

---

12. On utilise le même type de raisonnement utilisé pour la copule gaussienne pour démontrer ce résultat.

13. Introduites par Gesnest et Mackay.



- **Gumbel** :  $\varphi_{\theta \geq 1}(u) := (-\ln u)^\theta$
- **Franck** :  $\varphi_{\theta \geq 0}(u) := -\ln\left(\frac{e^{-\theta u}-1}{e^{-\theta}-1}\right)$
- **Clayton**  $\varphi_{\theta \neq 1}(u) := u^{-\theta} - 1$

La dépendance de queue à gauche de la copule **Gumbel** en dimension 2 est donnée par :

$$\begin{aligned}
 p^{min} &= \lim_{\alpha \rightarrow 0^+} \frac{C(u, u)}{u} \\
 &= \lim_{\alpha \rightarrow 0^+} \frac{\varphi_\theta^{-1}(2 \times \varphi_\theta(u))}{u} \\
 &\text{or } \phi^{-1}(u) = \exp\{-y^{\frac{1}{\theta}}\} \\
 &= \lim_{\alpha \rightarrow 0^+} \frac{\exp\{-(2 \times (-\ln u)^\theta)^{\frac{1}{\theta}}\}}{u} \\
 &= \lim_{\alpha \rightarrow 0^+} u^{2^{\frac{1}{\theta}}-1} \\
 &= 0 \\
 &\text{car } 2^{\frac{1}{\theta}} - 1 > 0 \Leftrightarrow \theta > 0 \text{ ce qui est vérifié par hypothèse.}
 \end{aligned}$$

Le coefficient de dépendance de queue à **droite** de la copule de **Gumbel** est déterminé par la **règle de l'Hôpital**

$$\begin{aligned}
 p^{max} &= \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u} \\
 &= \lim_{u \rightarrow 1^-} \frac{-2 + 2^{\frac{1}{\theta}} u^{2^{\frac{1}{\theta}}}}{-1} \\
 &= 2 - 2^{\frac{1}{\theta}}
 \end{aligned}$$

Si  $\theta > 1$  la copule de Gumbel a une dépendance de queue à droite. Si  $\theta = 1$ , la copule de Gumbel n'a pas de dépendance de queue.

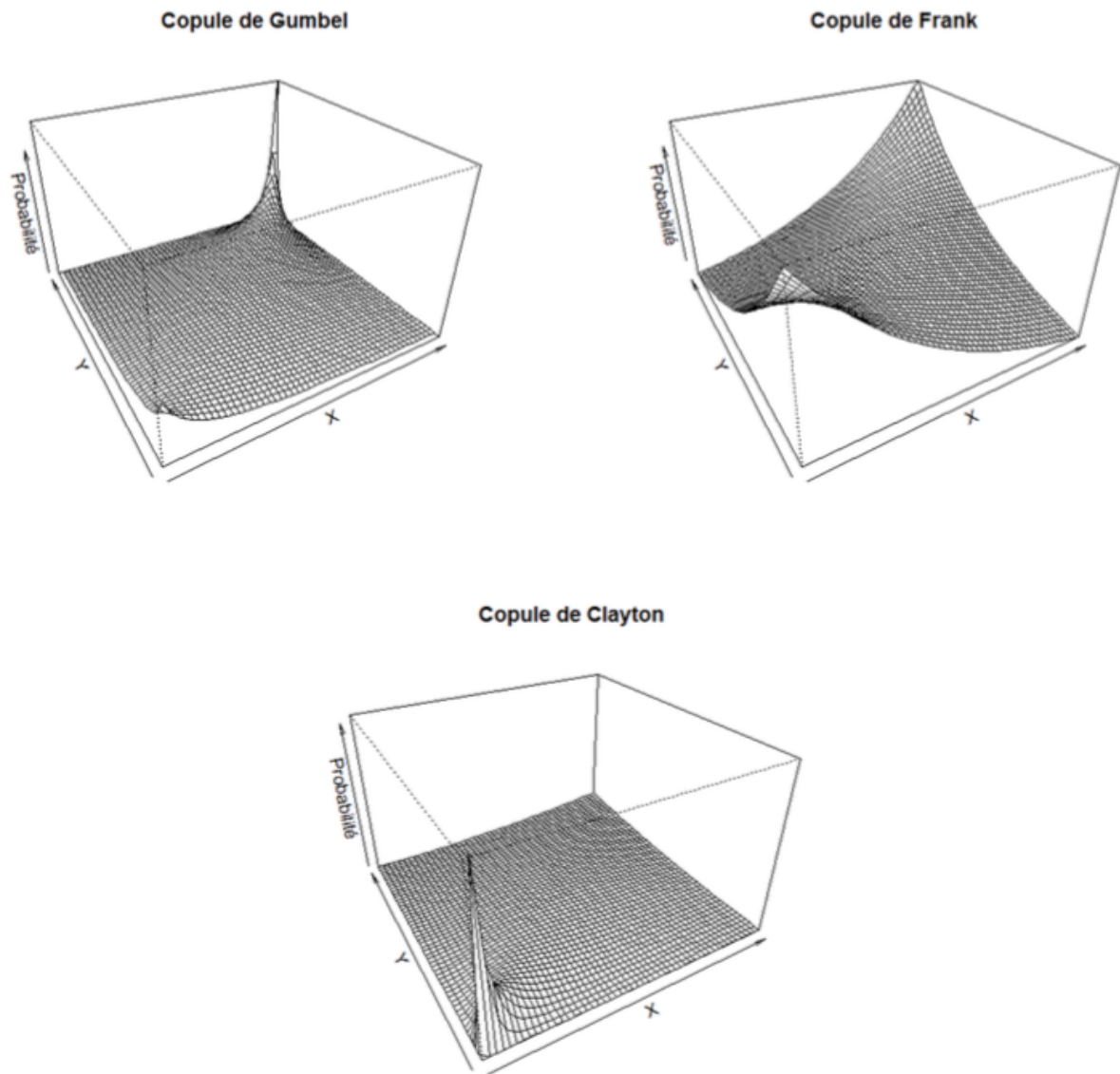


FIGURE A.1 – Densités de copules en dimension 2 pour un même paramètre  $\theta$

Ci-dessus le graphique de densités de copules archimédiennes permettant d'**illustrer** les structures de dépendance en fonction de la copule. On remarque que :

- la copule de **Gumbel** montre des concentrations aux niveaux des queues en particulier à droite.
- la copule de **Franck** montre des concentrations plus élevés dans les extrêmes mais de façon symétrique.

— la copule de **Clayton** montre des concentrations élevés au niveau de la queue à gauche.

Ainsi la famille de copule **archimédienne** permet de construire des densités jointes avec des structures de dépendance très **riches**.

En **assurance** ou en **finance**, la copule de **Gumbel** est intéressante pour modéliser des structures de dépendance des événements à **fortes** densités.

La copule de **Clayton** permet au contraire de la copule de Gumbel de modéliser des structures de de dépendance des événements de **faibles** intensités.

La copule de **Franck** permet de modéliser des structures de dépendances **symétriques** telles que les copules elliptiques de Gauss et de student.

Outre la richesse de structure de dépendance des copules archimédiennes, la **simulation** de ces copules **n'est pas forcément** compliquée.

Par exemple, pour la distribution de **Clayton** en dimension 2, il est possible d'inverser<sup>14</sup> la distribution **conditionnelle** par une formule **fermée** :

$$\begin{aligned} u_2 &= C_{2|1}(\bullet; u_1^\perp)^{-1}(u_2^\perp) \\ &= -\frac{1}{\theta} \ln \left( 1 + \ln \frac{u_2^\perp (\exp(-\theta) - 1)}{u_2^\perp + (1 - u_2^\perp) \exp(-\theta u_1^\perp)} \right) \end{aligned}$$

mais de façon **générale** pour les copules archimédiennes, on peut écrire la distribution conditionnelle d'une **en fonction** de son générateur et on obtient<sup>15</sup>

$$u_2 = \varphi^{-1} \left( \varphi^{(1)-1} \left( \varphi^{(1)} \left( \frac{u_1^\perp}{u_2^\perp} \right) \right) - \varphi(u_1^\perp) \right)$$

Nous avons ainsi dans cette section décrit les copules (semi)-paramétriques ainsi que des méthodes de simulations. Avant de pouvoir **opérationnellement** utiliser le concept de copule, il reste un point encore à décrire, il s'agit de l'étape d'**estimation**.

Cette étape **n'est** souvent **pas** très aisée. Néanmoins, présentons quelques méthodes **usuelles**. Nous nous attarderons pas sur cette étape car la méthode de simulation que nous présenterons par la suite s'appuie sur une copule **non-paramétrique** qui fera l'objet du chapitre suivant.

La méthode principale d'**estimation** est la fondée sur la **maximisation de la vraisemblance**<sup>16</sup> :

$$l(x_1, \dots, x_H; \Theta, B) := \sum_{k=1}^H \ln c_\theta \left( F_{\beta_1}(x_{1,1}), \dots, F_{\beta_d}(x_{d,k}) \right) + \sum_{k=1}^H \sum_{i=1}^d f_{\beta_i}(x_{i,k})$$

14. On sait écrire la distribution de la copule seulement comme une fonction relativement simple du paramètre  $\theta$ .

15. Cet algorithme est décrit dans l'article de C. Genest et R. MacKay [19] page 157.

16. On utilise donc l'expression de la **densité** d'une copule. *Two-stage parametric ML method* - proposé par **Shih** et **Louis** en 1985.

$$\forall i \quad \hat{\beta}_i = \operatorname{argmax}_{\theta} \sum_{k=1}^H f_{\beta_i}(x_{i,k})$$

puis on détermine l'estimateur du paramètre de copule :

$$\hat{\theta} = \operatorname{argmax}_{\beta_i} \sum_{k=1}^H \ln c_{\theta} \left( F_{\hat{\beta}_1}(x_{1,1}), \dots, F_{\hat{\beta}_d}(x_{d,k}) \right)$$

Pour certaines copules, par exemple la copule archimédienne de **Gumbel**, il est possible d'estimer le paramètre de la copule **simplement** en fonction de l'estimateur **empirique** de la corrélation de **Kendall** :

Le  $\tau$  de **Kendall** qui représente la différence entre la probabilité de concordance et de non-concordance :

$$\begin{aligned} \tau &= \mathbb{P}\{(X_1 - X_2)(Y_1 - Y_2) > 0\} - \mathbb{P}\{(X_1 - X_2)(Y_1 - Y_2) < 0\} = 2 \mathbb{P}\{(X_1 - X_2)(Y_1 - Y_2) > 0\} - 1 \\ &= 2 \times \left( \mathbb{P}\{X_1 > X_2, Y_1 > Y_2\} + \mathbb{P}\{X_1 < X_2, Y_1 < Y_2\} \right) - 1 \\ &= 2 \times \left( \int_{\mathbb{R}^2} 1 - F_{X_1}(x) - F_{Y_1}(y) + \mathbb{P}\{X_1 < x, Y_1 < y\} d\mathbb{P}(x, y) + \int_{\mathbb{R}^2} \mathbb{P}\{X_1 < x, Y_1 < y\} d\mathbb{P}(x, y) \right) - 1 \\ &= 2 \times \left( \int_{[0,1]^2} 1 - u_1 - u_2 + 2C(u_1, u_2) dC(u_1, u_2) \right) - 1 \\ &= 2 \times \left( 1 - \int_{[0,1]} u_1 du_1 - \int_{[0,1]} u_2 du_2 + 2 \int_{[0,1]^2} C(u_1, u_2) dC(u_1, u_2) \right) - 1 \\ &= 2 \times \left( 1 - \frac{1}{2} - \frac{1}{2} + 2 \int_{[0,1]^2} C(u_1, u_2) dC(u_1, u_2) \right) - 1 \\ &= 4 \int_{[0,1]^2} C(u_1, u_2) dC(u_1, u_2) - 1 \\ &= 4 \mathbb{E}[C(U_1, U_2)] - 1 \end{aligned}$$

Souvent cette espérance est difficile à calculer. Mais pour les copules archimédiennes on a le résultat <sup>17</sup> :

$$K(x) = \mathbb{P}\{C_{\theta}(U_1, U_2) < x\} = x - \frac{\varphi_{\theta}(x)}{\varphi'_{\theta}(x)}$$

On trouve donc simplement :

$$\mathbb{E}[C_{\theta}(U_1, U_2)] = \int_{[0,1]} xK'(x)dx = \frac{1}{2} \left( 1 - \frac{1}{2\theta} \right)$$

Le paramètre  $\theta$  de la copule de **Gumbel** est déterminé finalement par l'estimateur empirique <sup>18</sup> du  $\tau$  de **Kendall** :

$$\hat{\theta} = \frac{1}{1 - \hat{\tau}}$$

17. Fonction de **Kendall** - Cf. l'article de C. Genest et L-P Rivest.

18. Cf. page 52.

# Annexe B

## Tests statistiques

### B.1 Tests de significativité

#### B.1.0.1 Coefficient de Fisher - ou de détermination $R^2$

Le calcul du coefficient de détermination  $R^2$  du modèle (5.5) est **identique** à celui calculé pour la régression linéaire multiple :

$$\widehat{R}^2 := 1 - \frac{\sum_{i=1}^H \widehat{\epsilon}_t^2}{\sum_{i=1}^H (x_{t-i+1} - \bar{x})^2}$$

Cet indicateur indique le pourcentage de variance expliquée par le modèle. Il mesure la qualité d'ajustement.

Pour prendre en compte le **nombre** de variables de **retard** du modèle, on introduit le **coefficient de détermination ajusté** :

$$R_{Adj}^2 := 1 - \frac{H-1}{H-p} \frac{\sum_{i=1}^H \widehat{\epsilon}_t^2}{\sum_{i=1}^H (x_{t-i+1} - \bar{x})^2}$$

Ces coefficients varient entre 0 et 1. Le pouvoir **prédictif** du modèle est **fort** si  $\sum_{i=1}^H \widehat{\epsilon}_t$  tend vers 0.

#### Remarque B.1.1

Dans une régression **sans** constante le  $R^2$  n'est pas **interprétable** en terme de décomposition de la variance. On peut également remarquer que le dénominateur, variance totale, de ces indicateurs est **constant** quel que soit le modèle.

#### B.1.0.2 Test de Fisher

Le test de Fisher teste la significativité **globale** du modèle (5.5).

L'hypothèse **nulle** consiste à considérer  $\alpha_1 = 0 = \dots = \alpha_p = 0$ .

La statistique de Fisher est définie par :

$$F := \frac{\sum_{i=1}^H H_0 \widehat{\epsilon}_t^2 - \sum_{i=1}^H \widehat{\epsilon}_t^2}{\sum_{i=1}^H \widehat{\epsilon}_t^2}$$

où  ${}_{H_0}\hat{\epsilon}_t^2$  sont les résidus sous l'hypothèse  $H_0$ .

On montre que  $F \stackrel{\mathcal{L}}{\rightsquigarrow} F(p-1, H-p)$ , loi de Fisher-Snedecor de degrés de libertés  $p-1$  et  $H-p$ .

### B.1.0.3 Test de student

Le test de student<sup>1</sup> teste la significativité individuelle de chaque paramètres  $\alpha_i$ .

L'hypothèse nulle consiste à considérer  $\alpha_i = 0$  sans émettre d'hypothèses sur les autres  $\alpha_j$

La statistique de student est définie par :

$$\hat{T} := \frac{\hat{\alpha}_i}{\sqrt{\hat{V}(\hat{\alpha}_i)}}$$

On montre que sous l'hypothèse que le bruit blanc soit gaussien  $T \stackrel{\mathcal{L}}{\rightsquigarrow} t_{H-p}$ , loi de Student de degré de liberté  $H-p$ .

## B.2 Tests d'adéquation

### B.2.0.1 Test de Kolmogorov-Smirnov

Considérons la distribution empirique  $\hat{F}$  des résidus.

Le test de Kolmogorov-Smirnov permet de tester l'adéquation de la distribution empirique  $\hat{F}$  à une distribution théorique  $F^{th}$ .

L'hypothèse nulle est l'égalité des loi :  $\hat{F} = F^{th}$

La statistique de test de Kolmogorov-Smirnov est définie par :

$$D_H := \sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F^{th}|$$

Le test consiste à rejeter l'hypothèse de normalité pour des grandes valeurs  $D_H$ .

### B.2.0.2 Test de normalité de Shapiro-Wilk

Le test de Shapiro-Wilk teste l'adéquation de la distribution empirique à la loi normale en particulier pour des petits effectifs ( $n \leq 50$ ) :

La statistique de test est donnée<sup>2</sup> par :

$$W := \frac{\left( \sum_{i=1}^{\lfloor \frac{H}{2} \rfloor} a_i (\hat{\epsilon}_{(H-i+1:H)} - \hat{\epsilon}_{(i:H)}) \right)^2}{\sum_{i=1}^H (\hat{\epsilon}_i - \bar{\epsilon})^2}$$

1. Test de Wald dans le cas de la régression.

2.  $x_{(i)}$  désigne la statistique d'ordre Cf. page 165.

où les **constantes**  $a_i$  sont issues de la table de coefficient dite de Shapiro-Wilk.

L'hypothèse **nulle**  $H_0$  est la normalité des résidus.

Les valeurs de seuils critiques pour risque de première espèce  $1 - \alpha$  sont **tabulées**. Il y a rejet de l'hypothèse  $H_0$  pour une statistique  $W$  **faible**.

### B.2.0.3 Test de normalité de Jarque-Bera

C'est un test populaire qui vérifie également l'adéquation de la distribution empirique à la loi **normale** en se fondant sur les coefficients d'**asymétrie** et d'**aplatissement** de Pearson de la distribution des résidus et en les comparant avec les **valeurs de référence** de la loi normale :

$$\begin{aligned}\hat{\gamma} &:= \frac{\frac{1}{H} \sum_{i=1}^n (\epsilon_i - \tilde{\epsilon})^3}{\left(\frac{1}{H} \sum_{i=1}^n (\epsilon_i - \tilde{\epsilon})^2\right)^{\frac{3}{2}}} \\ \hat{\kappa} &:= \frac{\frac{1}{H} \sum_{i=1}^n (\epsilon_i - \tilde{\epsilon})^4}{\left(\frac{1}{H} \sum_{i=1}^n (\epsilon_i - \tilde{\epsilon})^2\right)^2}\end{aligned}$$

On montre que la loi du couple  $(\gamma, \kappa)$  tend vers une loi **normale bivariée**<sup>3</sup> :

$$\sqrt{H} \begin{pmatrix} \gamma \\ \kappa \end{pmatrix} \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}\left(\begin{pmatrix} 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 6 & 0 \\ 0 & 24 \end{pmatrix}\right)$$

L'hypothèse nulle  $H_0$  est la normalité.

La statistique de **Jarque-Bera** est définie par :

$$\hat{T} := H\left(\frac{\hat{\gamma}^2}{6} + \frac{(\hat{\kappa} - 3)^2}{24}\right)$$

On montre<sup>42</sup> que  $T \overset{\mathcal{L}}{\rightsquigarrow} \chi_2^2$ , loi du khi-deux à deux degré de libertés. L'hypothèse de normalité  $H_0$  pour un risque de premier espèce  $1 - \alpha$  sera **rejetée** si

$$\hat{T} > q_{1-\alpha}(\chi_2^2)$$

Ce test doit être plutôt utilisé lorsque les effectifs sont **élevés**. Il existe d'autres tests de normalité, (**Lilliefors**, **Anderson-Darling**, **d'Agostino**), on peut se référer par exemple au document de R. Rakotamalala [62] pour une description. Ces tests sont tous des variantes du test d'adéquation de Kolmogorov-Smirnov à l'instar de celui de **Jarque-Bera**.

3. Il existe une version plus précise pour la matrice de variance-covariance.

4. Par définition de la loi du khi-deux.

## B.3 Tests d'auto-corrélation sérielles des résidus

### B.3.0.1 Test de Durbin-Watson

Dans le cadre de régressions **sans** variables **retardées** de la variable **endogène** en tant que variables explicatives :

$$y_t = \alpha_0 + f(x_{1t}, \dots, x_{kt}; \alpha_1, \dots, \alpha_k) + \epsilon_t$$

l'éventuelle corrélation sérielle du processus  $\{\epsilon_t, t \in \mathbb{Z}\}$  est testé via le test populaire de Durbin-Watson.

Ce test consiste à tester la dépendance auto-régressive d'ordre 1 des résidus  $\epsilon_t$  :

$$\epsilon_t = \rho \epsilon_{t-1} + u_t \quad \text{u appelé innovation est un bruit blanc gaussien}$$

L'hypothèse nulle  $H_0$  consiste à considérer  $\rho = 0$ . La statistique de Durbin-Watson est définie par :

$$\widehat{DW} = \frac{\sum_{i=1}^{H-1} (\hat{\epsilon}_{t-i} - \hat{\epsilon}_{t-i-1})^2}{\sum_{i=1}^H \hat{\epsilon}_{t-i}^2}$$

où  $\hat{\epsilon}_{t-i}$  est le  $i^{\text{ème}}$  résidus **observés** avant  $t$ .

La statistique de  $\widehat{DB}$  qui prend des valeur entre 0 et 4. L'absence d'auto-corrélation sérielle est significative lorsque la statistique  $\widehat{DB}$  est **proche** de 2. Une valeur inférieure à 2 suggère une auto-corrélation positive. Une valeur supérieure à 2 suggère une auto-corrélation négative.

Mais pour obtenir des valeurs **critiques**, on se réfère à la **table** de Durbin-Watson. Néanmoins cette table fait apparaître des zones d'**indétermination**.

### B.3.0.2 Test de Breusch-Godfrey

Le test de Durbin-Watson n'est pas **pas** valide<sup>5</sup> en présence de variables **retardées** de la variable **endogène**<sup>6</sup> en tant que variable explicative c'est-à-dire l'hypothèse suivante doit être vérifiée :

$$\mathbb{C}orr(X_{t-s}, \epsilon_t) = 0$$

Or cette hypothèse n'est pas vérifiée par essence dans nos problématiques économétriques qui fait intervenir des modélisations auto-régressives de la variable endogène.

De plus le test de Durbin-Watson ne permet pas de tester des phénomènes d'auto-corrélation d'ordre **supérieur** à 1 i.e.

$$\epsilon_t = \rho_1 \epsilon_{t-1} + \dots + \rho_k \epsilon_{t-k} + u_t \quad \{u_t, t \in \mathbb{Z}\} \text{ est un bruit blanc}$$

Le test de Breusch-Godfrey<sup>7</sup> permet de contourner ses contraintes, il demeure **valide** même si la variable retardée est également un régresseur et permet de tester un ordre de corrélation sériel supérieur à 1. C'est donc le test que nous **privilégierons** par la suite.

5. Introduction d'un biais dans la statistique  $DW$ .

6. Appelé aussi variable dépendante.

7. C'est un test de type, test de multiplicateur de Lagrange - LM test.



Le test de **Breush-Godfrey** est construit à partir d'une régression **auxiliaire**

$$\widehat{\epsilon}_t = \rho_1 \widehat{\epsilon}_{t-1} + \dots + \rho_k \epsilon_{t-k} + f(x_{t-1}, \dots, x_{t-p}; \alpha_1, \dots, \alpha_p) + u_t \quad (\text{B.1})$$

L'hypothèse **nulle** est la nullité des coefficients  $\rho_i = 0$ . L'hypothèse alternative  $H_1$  est que les  $\epsilon_t$  suivent un processus  $AR(p)$  ou  $MA(p)$ .

La statistique de test de **Breush-Godfrey** est :

$$\widehat{LM} = H \widehat{R}^2$$

avec  $R^2$  le coefficient de **détermination** de la régression auxiliaire (B.1).

Sous l'hypothèse nulle,  $LM \stackrel{\mathcal{L}}{\rightsquigarrow} \chi_p^2$ , loi du **khi-deux** à  $p$  degrés de liberté.

L'hypothèse d'auto-corrélation des résidus sera rejetée si  $\widehat{LM} > q_{1-\alpha}(\chi_p^2)$  ou si la probabilité critique est petite.

Ce test n'est pas robuste en cas d'hétéroscédasticité conditionnelle des résidus.

### B.3.0.3 Tests de Box-Pierce et de Ljung-Box

Le test de **Box-Pierce** permet de tester l'**absence** d'auto-corrélation. On cherche à savoir si les résidus suivent un processus **bruit blanc** faible i.e.

- moyenne nulle
- volatilité constante
- pour tout  $k$ ,  $\text{Cov}(\epsilon_t, \epsilon_{t-k}) = 0$

L'hypothèse **nulle** est la nullité de toutes les auto-corrélations d'ordre  $k$  :  $\rho_1 = \dots = \rho_k = 0$ .

La statistique de **Box-Pierce** est définie par

$$\widehat{Q}_k := H \sum_{i=1}^k \widehat{\rho}_i^2 \quad \text{où } k \text{ est le nombre de retards.}$$

On montre que sous l'hypothèse **nulle**,  $Q_k \stackrel{\mathcal{L}}{\rightsquigarrow} \chi_k^2$ , loi du khi-deux à  $k$  degrés de liberté.

L'hypothèse nulle est **rejetée** si  $\widehat{Q}_k > q_{1-\alpha}(\chi_p^2)$  ou si la probabilité critique est petite.

Il existe une variante de ce test présentant de meilleure propriété asymptotique : le test de **Box-Pierce modifié**<sup>8</sup>

---

8. Appelé également test de **Ljung-Box**.

dont la statistique de test est donnée par :

$$\widehat{Q}_k := H(H+2) \sum_{h=1}^k \frac{\widehat{\rho}_h^2}{H-h} \quad \text{où } k \text{ est le nombre de retards.}$$

Ces tests sont appelés également **Q-tests** ou **tests portmanteau**<sup>9</sup>.

#### B.3.0.4 Test de nullité de la moyenne des résidus

Si le processus  $\{\epsilon_t, t \in \mathbb{Z}\}$  est un bruit blanc fort alors l'hypothèse **i.i.d** est **vérifiée** et le processus est de moyenne **nulle** et de variance  $\sigma_\epsilon$ . Donc d'après la loi forte des grands nombres :

$$\bar{\epsilon}_t := \frac{1}{H} \sum_{i=1}^H \widehat{\epsilon}_t \xrightarrow{H \rightarrow +\infty} 0$$

Par application du théorème centrale limite on montre que

$$\sqrt{H} \frac{\bar{\epsilon}_t}{\widehat{\sigma}_{\epsilon_t}} \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}(0, 1)$$

On peut tester la nullité des résidus en construisant l'intervalle de confiance d'un niveau de confiance  $1 - \alpha$  :

$$I_C(\bar{\epsilon}_t) = \left[ -\frac{\widehat{q}_{1-\alpha} \widehat{\sigma}_{\epsilon_t}}{\sqrt{H}}, \frac{\widehat{q}_{1-\alpha} \widehat{\sigma}_{\epsilon_t}}{\sqrt{H}} \right]$$

où  $q_{1-\alpha}$  est le quantile pour une niveau de confiance  $1 - \alpha$  de la loi normale centrée réduite.

## B.4 Tests d'hétéroscédasticité conditionnelle

### B.4.0.1 Test d'Engle

Le test d'Engle<sup>10</sup> cherche à valider l'hypothèse d'**hétéroscédasticité conditionnelle** des résidus.

L'hypothèse **nulle**  $H_0$  est l'**homoscédasticité** dans un modèle auto-régressif :

$$X_t := \Phi_{\alpha_0, \alpha_1}^{\text{AR}}(X_{t-1}) + \epsilon_t \quad \text{avec} \quad \epsilon_t | \mathcal{F}_t \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}(0, \sigma_\epsilon^2)$$

L'hypothèse **alternative**  $H_1$  est l'**hétéroscédasticité** conditionnelle :

$$X_t := \Phi_{\alpha_0, \alpha_1}^{\text{AR}}(X_{t-1}) + u_t \sqrt{\rho_0 + \rho_1 \epsilon_{t-1}^2} \quad u_t \text{ est un bruit blanc gaussien de variance unitaire}$$

c'est à dire<sup>11</sup>

$$\sigma_\epsilon^2 = \mathbb{E}[u_t^2 | u_{t-1}] = \rho_0 + \rho_1 u_{t-1}^2$$

9. Signifie en anglais littéralement mot valise, mélange.

10. C'est un test de type **multiplicateur de Lagrange**. ARCH – LM test en anglais.

11. On peut généraliser avec p retards sans difficulté.

Pour construire le test, on utilise donc une régression **auxiliaire** :

$$\widehat{\epsilon}_t^2 = \rho_0 + \rho_1 \widehat{\epsilon}_{t-1}^2 + u_t$$

La statistique de test d'Engle est définie via le coefficient de détermination de cette régression auxiliaire :

$$\widehat{LM} = H\widehat{R}^2$$

Sous l'hypothèse nulle  $H_0$ ,  $LM \stackrel{\mathcal{L}}{\rightsquigarrow} \chi_1^2$ , loi du khi-deux à 1 degré de liberté.

L'hypothèse d'hétéroscédasticité des résidus sera acceptée si

$$\widehat{LM} > q_{1-\alpha}(\chi_1^2)$$

ou si la probabilité critique est petite.

## B.5 Tests de stationnarité

### B.5.0.1 Test de Dickey-Fuller augmenté

Le test de Dickey-Fuller augmenté<sup>12</sup> permet de tester la stationnarité de séries temporelles de type **auto-régressive** avec une constante et une composante tendancielle :

$$A(L)X_t = \mu + \lambda t + \epsilon_t \tag{B.2}$$

où  $A(L)$  est un polynôme retard d'ordre  $p$  c'est-à-dire  $A(L) = 1 - \phi_1 L - \dots - \phi_p L^p$

Or on peut écrire<sup>13</sup>

$$A(L) = A(1) - \sum_{i=0}^{p-1} \phi'_i (L^i - L^{i+1})$$

avec

$$\begin{cases} \phi'_0 = \Phi(1) - 1 \\ \phi'_i = \phi'_{i-1} + \phi_i \end{cases}$$

En ré-arrangeant les termes de (B.2) on trouve :

$$\Delta X_t = \mu + \lambda t - A(1)X_{t-1} - \sum_{i=1}^{p-1} \phi'_i \Delta X_t + \epsilon_t \quad \text{avec} \quad \phi'_i = \phi_{i+1} + \dots + \phi_p$$

Par exemple, à l'ordre 3 (sans constante et tendance)

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \epsilon_t \\ X_t - X_{t-1} &= (\phi_1 - 1)X_{t-1} + (\phi_2 + \phi_3)X_{t-2} - \phi_3 \Delta X_{t-2} + \epsilon_t \\ \Delta X_t &= (\phi_1 + \phi_2 + \phi_3 - 1)X_{t-1} - (\phi_2 + \phi_3) \Delta X_{t-1} - \phi_3 \Delta X_{t-2} + \epsilon_t \end{aligned}$$

12. On note souvent test DFA.

13. Division des polynômes.

On peut donc écrire la dynamique (B.2) :

$$\Delta X_t = \mu + \lambda t - \Phi(1)X_{t-1} - \sum_{i=1}^{p-1} \phi'_i \Delta X_{t-i} + \epsilon_t$$

L'hypothèse **nulle** du test de Dickey-Fuller est  $\Phi(1) = 0$ . L'idée de l'équation de régression de Dickey-Fuller est donc d'obtenir des processus intégrés (stationnaires) si l'hypothèse nulle est vérifiée.

Prenons le cas particulier simple de la marche aléatoire, processus **non-stationnaire** :

$$X_t = \phi X_{t-1} + \epsilon_t$$

Le principe du test de Dickey-Fuller est donc de tester la présence d'une **racine unitaire**

$$\Phi(1) = 0 \Leftrightarrow 1 - \phi = 0 \Leftrightarrow \phi = 1.$$

Tester la présence d'une racine unitaire correspond **ainsi** à poser comme hypothèse **nulle**  $H_0$  la **non-stationnarité** du processus.

De prime abord, on souhaite tester la significativité du paramètre  $\phi$  et appliquer le **test de Wald**. La statistique de **student** est donnée par :

$$\hat{T} = \frac{\hat{\phi} - 1}{\sqrt{\text{Var}(\hat{\phi})}}$$

Or on montre que sous l'hypothèse nulle  $H_0$ , la statistique de **student**  $T$  **ne tend pas** vers une loi de **student** mais vers une loi **non-standard**<sup>14</sup> :

$$T \stackrel{\mathcal{L}}{\rightsquigarrow} \frac{1}{2} \frac{W(1)^2 - 1}{\sqrt{\int_0^1 W(s)ds}}$$

où  $W$  est un mouvement brownien.

Les seuils critiques de cette loi non-symétrique ont été **tabulée** par Dickey-Fuller.

Nous avons cherché la loi asymptotique dans l'exemple simple d'une marche aléatoire mais pour les autres cas, lorsqu'on introduit la constante et/ou la tendance temporelle (B.2) page 155, on montre que la loi asymptotique de la statistique de student **dépend** du modèle **considéré** et donc avec des seuils critiques différents.

Le test de Dickey-Fuller **augmenté** est donc plutôt un **ensemble** de tests de type **student** en fonction du modèle considéré (avec ou sans constante et/ou avec ou sans composante temporelle). On doit donc mettre en place une **stratégie** de tests<sup>15</sup>.

Le test dit de Dickey-Fuller est un cas **particulier** du test de Dickey-Fuller **augmenté** avec  $p = 1$ .

14. Donner une référence.

15. Nous utilisons la fonction "adf.test" du package "tseries" du logiciel R Cf. [78].

### B.5.0.2 Test KPSS

Le test KPSS est un test de stationnarité dont l'hypothèse **nulle** est la **stationnarité** contrairement au test de Dickey-Fuller.

### B.5.0.3 Test de Chow

Le test de Chow permet de tester l'homogénéité de deux sous-population i.e dans le cadre de série temporelle de tester l'existence d'une **rupture temporelle** éventuelle dans la série.

On **choisit**<sup>16</sup> deux sous-périodes historique de  $(x_{t-H}, \dots, x_{t-H'})$  et  $(x_{t-H'+1}, \dots, x_{t-1})$ . On suppose que sur la première partie de l'historique la dynamique<sup>16</sup> est

$$Y_t^1 = \alpha_0^1 + f(X_{t-p}^1, \dots, X_{t-1}^1; \alpha_1^1, \dots, \alpha_p^1) + \epsilon_t^1$$

et sur la deuxième partie de l'historique, la dynamique<sup>17</sup> est :

$$Y_t^2 = \alpha_0^2 + f(X_{t-p}^2, \dots, X_{t-1}^2; \alpha_1^2, \dots, \alpha_p^2) + \epsilon_t^2$$

L'hypothèse **nulle** consiste à supposer  $\alpha_i^1 = \alpha_i^2$ . Soit  $SCR_C$  la somme des carrés des résidus dans le modèle contraint i.e. sous l'hypothèse nulle et  $SCR$  la somme des carrés des résidus dans le modèle sans contraintes.

La statistique de test est une statistique de Fisher :

$$F_{\text{Chow}} := \frac{SCR_C - SCR}{SCR} \frac{H - 2(p + 1)}{p + 1}$$

Cette statistique  $F_{\text{Chow}} \stackrel{\mathcal{L}}{\sim} \frac{1}{p+1} \chi_{p+1}^2$ , loi de khi-deux de degrés de liberté  $p + 1$ .

L'hypothèse sera **rejetée** pour des valeur **élevé** de  $F_{\text{Chow}}$ .

## B.6 Critère de comparaison de modèles

---

16. La difficulté est dans ce choix.

17. Ici on suppose que les variables explicatives ne sont pas des variables retardées de la variable dépendante.

## Annexe C

# Estimation paramétrique

### C.1 Moindres carrés non-linéaire

Dans ce paragraphe nous nous intéressons aux modèles de régression de la forme :

$$\forall t \in \mathbb{Z} \quad y_t = f(\underline{x}_{t-1}; \underline{\theta}) + \epsilon_t \quad (\text{C.1})$$

où

- $\underline{x}_t$  un vecteur
- $\underline{\theta}$  un vecteur de  $\mathbb{R}^k$
- $(\epsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc

#### C.1.0.1 Non-linéarité dans les variables

Le modèle de régression (C.1) est **non-linéaire** dans les **variables** si

$$\frac{\partial y_t}{\partial \theta_k} = g(\underline{x}_{t-1})$$

i.e. la dérivée partielle **ne dépend pas** des paramètres  $\theta_k$ .

On dit alors que le modèle est **linéaire** dans les *paramètres*. Avec des redéfinitions **appropriées** des régresseurs  $x_{t-1}$  on peut *ré-écrire* le modèle (C.1) sous la forme

$$\underline{y} = X\underline{\beta} + \epsilon \quad (\text{C.2})$$

Dans la matrice  $X$  de (C.2), les fonctions **peuvent être** non-linéaires dans les variables explicatives.

On peut donc calculer l'estimateur MCO *habituel* des paramètres  $\underline{\theta}$  donné par

$$\hat{\underline{\theta}} = (X^T X)^{-1} X^T \underline{y}$$

### C.1.0.2 Non-linéarité dans les paramètres

Le modèle de régression (C.1) est **non-linéaire** dans les **paramètres** si

$$\frac{\partial y_t}{\partial \theta_k} = h(\underline{x}_{t-1}, \underline{\theta})$$

i.e. la dérivée partielle **ne dépend pas** des paramètres  $\theta_k$ .

Dans ce cas, il n'est **plus** possible d'écrire le modèle sous la forme  $\underline{y} = X\underline{\beta} + \epsilon$ . Néanmoins il est toujours possible de minimiser la somme des erreurs au carré du modèle :

$$\min_{\underline{\theta}} \frac{1}{H-1} \sum_{i=1}^{H-1} (y_{t-i} - f(\underline{x}_{t-i-1}; \underline{\theta}))^2 = \min_{\underline{\theta}} \frac{1}{H-1} (\underline{y} - \underline{f}(\underline{x}_{t-1}, \dots, \underline{x}_{t-H}; \underline{\theta}))^T (\underline{y} - \underline{f}(\underline{x}_{t-1}, \dots, \underline{x}_{t-H}; \underline{\theta}))$$

mais il n'y a plus d'expression algébrique *simple*. Il faut écrire les conditions de premier ordre :

$$-\frac{2}{H-1} \nabla_{\theta} f(\underline{x}_{t-1}, \dots, \underline{x}_{t-H}; \underline{\theta}) (\underline{y} - \underline{f}(\underline{x}_{t-1}, \dots, \underline{x}_{t-H}; \underline{\theta})) = 0$$

où  $\nabla_{\theta} f(\underline{x}_{t-1}, \dots, \underline{x}_{t-H}; \underline{\theta})$  est le **gradient**, matrice de  $\mathcal{M}_{k,H}$

$$\nabla_{\theta} f(\underline{x}_{t-1}, \dots, \underline{x}_{t-H}; \underline{\theta}) := \left( \nabla_{\theta} f(\underline{x}_{t-1}; \underline{\theta}), \dots, \nabla_{\theta} f(\underline{x}_{t-H}; \underline{\theta}) \right)$$

### C.1.0.3 Moindres carrés conditionnels - Estimation des processus SETAR

Dans le cadre des processus SETAR nous remarquons que **conditionnellement** à l'évènement  $X_{t-1} \geq c$  le processus devient **linéaire** en les paramètres :

$$\mathbb{E}[X_t | X_{t-1} \leq c] = (\alpha_{11}, \alpha_{21}) \tilde{X}_{t-1}^1 + (\alpha_{21}, \alpha_{22}) \tilde{X}_{t-1}^2$$

où les vecteurs  $X_{t-1}^i$  sont définis par

$$\begin{aligned} \tilde{X}_{t-1}^1 &:= \left( \mathbf{1}_{\{X_{t-1} \leq c\}}, X_{t-1} \mathbf{1}_{\{X_{t-1} \leq c\}} \right)^T \\ \tilde{X}_{t-1}^2 &:= \left( \mathbf{1}_{\{X_{t-1} > c\}}, X_{t-1} \mathbf{1}_{\{X_{t-1} > c\}} \right)^T \end{aligned}$$

Si l'on observe l'échantillon  $(x_{t-H}, \dots, x_{t-1})$  du n-uplet  $(X_{t-H}, \dots, X_{t-1})$ , on obtient les paramètres  $\alpha_{i,j}$  **conditionnellement à c** en minimisant l'erreur quadratique :

$$Q(c|x_{t-H}, \dots, x_{t-1}) := \sum_{i=2}^H (x_{t-i+1} - \mathbb{E}[X_{t-i+1} | X_{t-i} \leq c, X_{t-i} = x_{t-i}])^2$$

On obtient  $\hat{\alpha}_{i,j}(c)$ , l'estimateur conditionnel des moindres carrés de  $\alpha_{i,j}$

$$\hat{\alpha}_{i,j}(c) = (\bar{X}^T \bar{X})^{-1} \bar{X} X$$

C'est l'estimateur des moindres carrés ordinaires avec la matrice  $\bar{X}$  définie par

$$\bar{X} := \begin{pmatrix} \mathbf{1}_{\{x_{t-H} \leq c\}} & x_{t-H} \mathbf{1}_{\{x_{t-H} \leq c\}} & \mathbf{1}_{\{x_{t-H} > c\}} & x_{t-H} \mathbf{1}_{\{x_{t-H} > c\}} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{1}_{\{x_{t-2} \leq c\}} & x_{t-2} \mathbf{1}_{\{x_{t-2} \leq c\}} & \mathbf{1}_{\{x_{t-2} > c\}} & x_{t-2} \mathbf{1}_{\{x_{t-2} > c\}} \end{pmatrix}$$

## C.2 Quasi-Maximum de vraisemblance

### C.2.0.1 Vraisemblance

On observe<sup>1</sup> l'échantillon  $(x_{t-H}, \dots, x_{t-1})$  du n-uplet  $(X_{t-H}, \dots, X_{t-1})$ . On suppose que ce n-uplet est généré par un processus dynamique de série temporelle paramétrique. Le modèle est paramétrique de paramètre  $\theta$  de l'espace  $\Theta$  :

La vraisemblance de l'échantillon est donnée par la densité **jointe** :

$$L(\theta; x_{t-H}, \dots, x_{t-1}) = f_{X_{t-H}, \dots, X_{t-1}}(x_{t-H}, \dots, x_{t-1}; \theta)$$

or par définition de la loi conditionnelle on a :

$$\begin{aligned} f_{X_{t-H}, \dots, X_{t-1}}(x_{t-H}, \dots, x_{t-1}; \theta) &= f_{X_{t-H+1}, \dots, X_{t-1} | X_{t-H}}(x_{t-H+1}, \dots, x_{t-1} | x_{t-H}; \theta) f_{X_{t-H}}(x_{t-H}; \theta) \\ &= f_{X_{t-H}}(x_{t-H}) \prod_{i=2}^H f_{X_{t-i+1} | X_{t-H}, \dots, X_{t-i}}(x_{t-i+1} | x_{t-H}, \dots, x_{t-i}; \theta) \end{aligned}$$

En allégeant la notation<sup>2</sup> la log-vraisemblance du n-échantillon  $\underline{x}^H := (x_{t-H}, \dots, x_{t-1})$  devient

$$\begin{aligned} l(\theta; \underline{x}^H) &:= \ln f(x_{t-H}) + \sum_{i=2}^H \ln f(x_{t-i+1} | x_{t-H}, \dots, x_{t-i}; \theta) \\ &= \ln f(x_{t-H}) + \sum_{i=2}^p \ln f(x_{t-i+1} | x_{t-H}, \dots, x_{t-i}; \theta) + \sum_{i=p+1}^H \ln f(x_{t-i+1} | x_{t-H}, \dots, x_{t-i}; \theta) \end{aligned}$$

Si la partie de la vraisemblance qui correspond aux contributions des conditions initiales est **constante**, on parle alors de vraisemblance **conditionnelle** sinon on parlera de vraisemblance **non conditionnelle**.

On dit que  $\hat{\theta}^{\text{MV}}$  est un estimateur du **maximum de vraisemblance** si

$$\forall \theta \in \Theta \quad l(\hat{\theta}^{\text{MV}}; \underline{X}^H) \geq l(\theta; \underline{X}^H)$$

$l(\hat{\theta}^{\text{MV}}; \underline{x}^H)$  est une **estimation** du maximum de vraisemblance en  $\hat{\theta}^{\text{MV}}$ <sup>3</sup>.

#### Proposition C.2.1 (Equations de vraisemblance)

Si la fonction de vraisemblance **atteint** un maximum alors la log-vraisemblance doit satisfaire les conditions de premier ordre, **équations de vraisemblance** :

$$\nabla l(\theta; \underline{x}^H) = \nabla \ln f(x_{t-H}; \theta) + \sum_{i=2}^H \nabla \ln f(x_{t-i+1} | x_{t-H}, \dots, x_{t-i}; \theta) = \underbrace{(0, \dots, 0)}_{\dim(\Theta) \text{ fois}}$$

où  $\nabla$  est le **gradient**, vecteur **ligne**  $\nabla g(x) = (\partial_{x_1} g(x), \dots, \partial_{x_n} g(x))$ .

1.  $H$  est la longueur de l'historique.
2.  $p$  est le nombre de conditions initiales du processus générateur.
3. Il y a existence de  $\hat{\theta}^{\text{MV}}$  si  $l$  est continue sur  $\Theta$ , un espace compact.



**Proposition C.2.2 (Concentration de la vraisemblance)**

La recherche des solutions des équations de vraisemblance peut être **complexe**. Dans certains cas, il est possible de trouver la solution au programme d'optimisation en plusieurs étapes **successives**. Si le paramètre  $\theta$  se décompose en  $(\theta_1, \theta_2)$  on opère de la façon suivante :

1. à  $\theta_2$  **fixé**, on résout la première équation de vraisemblance  $\partial_{\theta_1} L(\theta_1; \theta_2, x^H) = 0$ . La solution dépend alors de  $\theta_2$ . On dit que la vraisemblance est **concentrée** par rapport à  $\theta_1$ .
2. On résout ensuite l'équation de vraisemblance  $\partial_{\theta_2} L(\theta_1(\theta_2), \theta_2; x^H) = 0$ . pour déterminer  $\hat{\theta}_2^{\text{MV}}$
3. En utilisant la relation  $\theta_1(\theta_2^{\text{MV}})$  on détermine enfin  $\hat{\theta}_1^{\text{MV}}$ .

Ce résultat s'appuie sur le **théorème de l'enveloppe**. Considérons le problème d'optimisation

$$\tilde{l}(\theta_2) := \operatorname{argmax}_{\theta_1} l(\theta_1; \theta_2)$$

Soit  $\hat{\theta}_1(\theta_2)$  tel que  $\hat{\theta}_1(\theta_2) = \operatorname{argmax}_{\theta_1} l(\theta_1; \theta_2)$  alors  $\tilde{l}(\cdot)$  est **différentiable**

$$\frac{\partial \tilde{l}(\hat{\theta}_2)}{\partial \theta_2} = \frac{\partial l(\hat{\theta}_1(\hat{\theta}_2), \hat{\theta}_2)}{\partial \theta_2}$$

Les conditions de premier ordre se simplifient alors. En effet la différentielle totale est donnée par :

$$dl(\theta_1, \theta_2) = \frac{\partial l(\theta_1, \theta_2)}{\partial \theta_1} d\theta_1 + \frac{\partial l(\theta_1(\theta_2), \theta_2)}{\partial \theta_2} d\theta_2 = \frac{\partial l(\theta_1, \theta_2)}{\partial \theta_1} d\theta_1$$

Du fait de la dépendance de l'estimateur  $\theta_1$  au n-échantillon  $x^H$ , la log-vraisemblance concentrée ne peut plus s'exprimer comme la somme des contributions de chaque observation. En effet

$$l(\theta_1, \theta_2; \underline{x}^H) = \sum_{i=1}^H \ln f(x_{t-i+1} | x_{t-H} \cdots, x_{t-i}; \theta_1(\theta_2; x^H), \theta_2)$$

**C.2.0.2 Estimation d'un modèle AR(1)**

**C.2.0.3 Pseudo-log-vraisemblance d'un modèle AR(1) – ARCH(1)**

On appelle **pseudo-maximum**<sup>4</sup> de vraisemblance d'un échantillon sa vraisemblance lorsqu'on suppose en **plus** que le bruit blanc est **gaussien**. En posant  $h_t^2 := \alpha_0 + \alpha_1 \epsilon_{t-1}^2$ , la loi conditionnelle est gaussienne :

$$X_t | \mathcal{F}_{t-1} \stackrel{\mathcal{L}}{\sim} \mathcal{N}(\Phi^{\text{AR}}(X_{t-1}), h_t^2)$$

La pseudo-log-vraisemblance conditionnelle s'écrit donc

4. On parle aussi de quasi-maximum de vraisemblance.

Ici il s'agit de processus autoregressifs d'ordre  $p = 1$ .

$$\begin{aligned}
l(\mu, \rho, \alpha_0, \alpha_1; \underline{x}^H) &= \sum_{i=p+1}^H \ln f(x_{t-i+1} | x_{t-H}, \dots, x_{t-i}; \mu, \rho, \alpha_0, \alpha_1) \\
&= \sum_{i=2}^H \ln \frac{1}{\sqrt{2\pi} h_{t-i+1}} e^{-\frac{1}{2} \frac{(x_{t-i+1} - \Phi_{\mu, \rho}^{\text{AR}}(x_{t-i}))^2}{h_{t-i+1}^2}} \\
&= -\frac{H-1}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=2}^H \ln h_{t-i+1}^2 - \frac{1}{2} \sum_{i=2}^H \frac{\epsilon_{t-i+1}^2}{h_{t-i+1}^2}
\end{aligned}$$

L'estimateur de pseudo-maximum de vraisemblance est *solution* du problème

$$(\hat{\mu}, \hat{\rho}, \hat{\alpha}_0, \hat{\alpha}_1) = \underset{\mu, \rho, \alpha_0, \alpha_1}{\operatorname{argmin}} \sum_{i=2}^H \left( \ln h_{t-i+1}^2 + \frac{\epsilon_{t-i+1}^2}{h_{t-i+1}^2} \right)$$

On résoud à partir des conditions de premiers ordre.

L'ensemble de ces relations constituent un systèmes d'équations non linéaires par rapport aux paramètres et doivent donc être résolues par des méthodes numériques.

**Proposition C.2.3 (Approche d'Engle - Algorithme du score)**

Si les termes non-diagonaux de la matrice d'information, matrice bloc, sont nuls alors on peut procéder à une estimation séparée des paramètres  $(\mu, \rho)$  et  $(\alpha_0, \alpha_1)$ .

**C.2.0.4 Pseudo-log-vraisemblance d'un modèle Soft – EXPAR(1)**

Le pseudo-maximum de vraisemblance d'un échantillon d'un processus Soft – ExpAR(1) est donné par :

$$\begin{aligned}
l(\mu, \rho_1, \rho_2, \sigma_\epsilon; \underline{x}^H) &= \sum_{i=2}^H \ln f(x_{t-i+1} | x_{t-H}, \dots, x_{t-i}; \mu, \rho_1, \rho_2) \\
&= \sum_{i=2}^H \ln \frac{1}{\sqrt{2\pi} \sigma_\epsilon} e^{-\frac{1}{2} \frac{(x_{t-i+1} - \Phi_{\mu, \rho_1, \rho_2}^{\text{Soft-ExpAR}}(x_{t-i}))^2}{\sigma_\epsilon^2}} \\
&= -\frac{H-1}{2} \ln(2\pi) - (H-1) \ln \sigma_\epsilon - \frac{1}{2\sigma_\epsilon^2} \sum_{i=2}^H \epsilon_{t-i+1}^2
\end{aligned}$$

L'estimateur de pseudo-maximum de vraisemblance est solution du programme d'optimisation :

$$(\hat{\mu}, \hat{\rho}_1, \hat{\rho}_2, \hat{\sigma}_\epsilon) = \underset{\mu, \rho_1, \rho_2, \sigma_\epsilon}{\operatorname{argmin}} (H-1) \ln \sigma_\epsilon + \frac{1}{2\sigma_\epsilon^2} \sum_{i=2}^H \epsilon_{t-i+1}^2$$

La première étape de la minimisation consiste à **concentrer** la pseudo-vraisemblance par rapport au paramètre  $\sigma_\epsilon$  et en écrivant l'équation de vraisemblance

$$\begin{aligned} \frac{\partial}{\partial \sigma_\epsilon} (H-1) \ln \sigma_\epsilon - \frac{1}{2\sigma_\epsilon^2} \sum_{i=2}^H \epsilon_{t-i+1}^2 &= 0 \\ \Leftrightarrow \frac{H-1}{\sigma_\epsilon} + \frac{1}{\sigma_\epsilon^3} \sum_{i=2}^H \epsilon_{t-i+1}^2 &= 0 \end{aligned}$$

La résolution de cette équation conduit à

$$\hat{\sigma}_\epsilon(\mu, \rho_1, \rho_2) = \left( \frac{1}{H-1} \sum_{i=2}^H \epsilon_{t-i+1}^2 \right)^{\frac{1}{2}}$$

La fonction de vraisemblance **concentrée** est donc donnée par :

$$l^c(\mu, \rho_1, \rho_2; \underline{x}^H) := -\frac{H-1}{2} \ln(2\pi) - \frac{(H-1)}{2} \ln \frac{1}{H-1} \sum_{i=2}^H \epsilon_{t-i+1}^2 - \frac{H-1}{2}$$

Finalement, maximiser de pseudo-vraisemblance concentrée est équivalent à minimiser la **somme des carrés des résidus** pour déterminer les estimateurs  $\hat{\mu}, \hat{\rho}_1, \hat{\rho}_2$  :

$$\min \frac{1}{H-1} \sum_{i=2}^H \epsilon_{t-i+1}^2$$

Il s'agit d'un modèle avec dépendance **linéaire** en les paramètres de type :  $x_{t+1} = \sum_{k=1}^n \theta_k \phi_k(x_t)$  Le problème consiste donc à minimiser la *distance euclidienne*

$$\min_{\theta=(\mu, \rho_1, \rho_2)} \sum_{i=2}^H (y_i - \tilde{x}_i \theta)^2 \quad (\text{C.3})$$

où

—  $y_i$  est le  $i^{\text{ème}}$  élément du vecteur  $\underline{y} := \begin{pmatrix} x_{t-H+1} \\ \vdots \\ x_{t-1} \end{pmatrix}$

—  $\tilde{x}_i$  est le  $i^{\text{ème}}$  ligne (vecteur ligne) de la matrice  $\tilde{\underline{x}} := \begin{pmatrix} 1 & x_{t-H} & x_{t-H} e^{-x_{t-H}^2} \\ \vdots & \vdots & \vdots \\ 1 & x_{t-2} & x_{t-2} e^{-x_{t-2}^2} \end{pmatrix}$

—  $\underline{\theta}$  le vecteur de paramètres  $\begin{pmatrix} \mu \\ \rho_1 \\ \rho_2 \end{pmatrix}$

Le problème (C.3) s'écrit en écriture *matricielle*

$$\min_{\underline{\theta}} (\underline{y} - \underline{\tilde{x}} \underline{\theta})^T (\underline{y} - \underline{\tilde{x}} \underline{\theta})$$

$\underline{\hat{\theta}}$  est déterminée par les **équations normales** (condition de premier ordre)

$$\underline{\tilde{x}}^T (\underline{y} - \underline{\tilde{x}}^T \underline{\hat{\theta}}) = 0$$

On obtient la formule *standard*  $\underline{\hat{\theta}} = (\underline{\tilde{x}}^T \underline{\tilde{x}})^{-1} \underline{y}$ .

Enfin, la variance s'exprime en fonction des paramètres estimés  $\underline{\hat{\theta}}$  :

$$\hat{\sigma}_\epsilon = \left( \frac{1}{H-1} \sum_{i=2}^H (x_{t-i+1} - \hat{\mu} - \hat{\rho}_1 - \hat{\rho}_2 x_{t-i} e^{x_{t-i}^2})^2 \right)^{\frac{1}{2}}$$

### C.2.0.5 Pseudo-estimateur MV d'un modèle ExpAR(1) – ARCH(1)

Pour déterminer le pseudo-estimateur du maximum de vraisemblance d'un processus on doit résoudre les équations de vraisemblance concentrées par rapport aux paramètres de variance conditionnelle :

$$\left\{ \begin{array}{l} \frac{\partial}{\partial \alpha_i} l(\mu, \rho, \alpha_0, \alpha_1; \underline{x}^H) = \sum_{i=1}^H \frac{1}{2h_{t-i+1}} \left( \frac{\epsilon_{t-i+1}^2}{2h_{t-i+1}^2} - 1 \right) \frac{\partial h_{t-i+1}^2}{\partial \alpha_i} = 0 \quad \forall i = 0, 1 \\ \frac{\partial}{\partial \beta} l(\beta := (\mu, \rho), \hat{\alpha}_0, \hat{\alpha}_1; \underline{x}^H) = -\frac{1}{2} \sum_{i=1}^H \frac{1}{h_{t-i+1}^2} \frac{\partial h_{t-i+1}^2}{\partial \beta} - \sum_{i=1}^H \frac{1}{h_{t-i+1}^2} \epsilon_{t-i+1} \frac{\partial \epsilon_{t-i+1}}{\partial \beta} + \frac{1}{2} \sum_{i=1}^H \epsilon_{t-i+1}^2 \frac{1}{h_{t-i+1}^4} \frac{\partial h_{t-i+1}^2}{\partial \beta} = 0 \\ \forall \beta = \mu, \rho \end{array} \right.$$

Les dérivées partielles dépendent de la fonction moyenne conditionnelle  $\Phi$ . Dans le cas du modèle ExpAR(1) – ARCH(1) on a

$$\frac{\partial h_{t-i+1}^2}{\partial \beta} = \alpha_1 \epsilon_{t-i+1} \frac{\partial \epsilon_{t-i+1}}{\partial \beta} = \alpha_1 \epsilon_{t-i+1} \partial_\beta \Phi(x_{t-i+1}) \quad \forall \beta = \mu, \rho$$

avec

$$\left\{ \begin{array}{l} \partial_\mu \Phi(x_{t-i+1}) = 1 \\ \partial_{\rho_1} \Phi(x_{t-i+1}) = x_{t-i+1} \\ \partial_{\rho_2} \Phi(x_{t-i+1}) = x_{t-i+1} e^{-x_{t-i+1}^2} \end{array} \right.$$

## Annexe D

# Statistiques d'ordre

Soient  $n$  observations  $(x_1, \dots, x_n)$  d'un  $n$ -échantillon  $(X_1, \dots, X_n)$ , vecteur *aléatoire* de variables *i.i.d.*

On note  $(x_{(1)}, \dots, x_{(n)})$  le vecteur des observations *ordonnées* par ordre **croissant**.

La **statistique d'ordre** associée à l'échantillon  $(X_1, \dots, X_n)$  est le vecteur  $(X_{(1:n)}, \dots, X_{(n)})$ .

$X_{(i)}$  est appelé la  $i^{\text{ème}}$  **statistique d'ordre**<sup>1</sup>.

Par définition, on a

$$— x_{(1)} = \min(x_1, \dots, x_n)$$

$$— x_{(n)} = \max(x_1, \dots, x_n)$$

L'ordre des observations est indiqué par les **rangs** des observations  $r_i$ .

La **statistique de rang** associée l'échantillon  $(X_1, \dots, X_n)$  est le vecteur  $(R_1, \dots, R_n)$  :

$$R_i = 1 + \sum_{j=1}^n 1_{\{X_j < X_i\}}$$

Le rang  $R_i$  de la  $i^{\text{ème}}$  observation  $X_i$  est appelé la  $i^{\text{ème}}$  **statistique de rang**.

---

1. On note également  $X_n^*$ . Lorsqu'on souhaite souligner la taille de l'échantillon, on note souvent la statistique d'ordre  $X_{(n)} := X_{(n:n)}$

On peut remarquer que s'il n'y pas d'ex æquo on a la relation suivante

$$\begin{aligned}r_i &= 1 + \sum_{j=1}^n 1_{\{x_j < x_i\}} \\ &= \sum_{j=1}^n 1_{\{x_j \leq x_i\}} \\ &= nF_n(x_i)\end{aligned}$$

En cas d'ex æquo on a la relation suivante :

$$F_n(x_i) = \frac{1}{n}(r_i + \max\{\#\{\text{ex æquo entre } i \text{ et } i+1\} - 1, 0\})$$

Néanmoins il est peu probable d'avoir des ex æquo.

Soit  $p \in ]0, 1[$ . La statistique d'ordre  $X([\!np] + 1)$  s'appelle **quantile empirique** d'ordre  $p$  du  $n$ -échantillon  $(X_1, \dots, X_n)$ .

Si la fonction de distribution  $F_X$  est continue et strictement croissante alors

$$X([\!np] + 1) \xrightarrow[p.s.]{} q_p$$

$X([\!np] + 1)$  est donc une estimation de  $q_p$  d'autant meilleure que la taille de l'échantillon  $n$  est grande.

## Annexe E

# Distribution de Pareto généralisée

### E.1 Graphique de l'espérance excédentaire

Soit  $(x_1, \dots, x_n)$  l'observation du n-échantillon  $(X_1, \dots, X_n)$ .

Le **graphe** de la fonction des excès au delà d'un seuil<sup>1</sup>  $u$  est défini par le graphe des **points**

$$\left\{ x_{(1:n)} \leq u \leq x_{(n:n)} < x_F : (u, e_n(u)) \right\}$$

où

$$e_n(u) := \frac{1}{N_u} \sum_{i=1}^n (x_i - u)^+$$

$e_n(u)$  est la moyenne **empirique** des excès au delà du seuil  $u$  :

$$e(u) = \mathbb{E}[X - u | X > u]$$

$N_u$  étant le **nombre** d'excès par rapport au seuil  $u$  :

$$N_u = \sum_{i=1}^n 1_{\{x_i > u\}}$$

Les lois GPD sont **stables** par l'opération de **seuil** c'est-à-dire :

$$\text{Si } X \overset{\mathcal{L}}{\rightsquigarrow} \text{GPD}(\xi, \sigma) \text{ alors } \forall u > u_0 \quad X - u | X > u \overset{\mathcal{L}}{\rightsquigarrow} \text{GPD}(\xi, \sigma + \xi(u - u_0))$$

---

1. Mean-Excess plot - ou Mean Residual Life Plot

En effet

$$\begin{aligned}
\mathbb{P}\{X - u \leq y \mid X > u\} &= \mathbb{P}\{X - u_0 \leq y + u - u_0 \mid X - u_0 > u - u_0\} \\
&= \frac{\mathbb{P}\{X - u_0 \leq y + u - u_0, X - u_0 > u - u_0\}}{\mathbb{P}\{X - u_0 \geq u - u_0\}} \\
&= \frac{F_{u_0}(y + u - u_0) - F_{u_0}(u - u_0)}{1 - F_{u_0}(u - u_0)} \\
&\sim \frac{H_{\xi, \sigma(u_0)}(y + u) - H_{\xi, \sigma(u_0)}(u)}{1 - H_{\xi, \sigma(u_0)}(u)} \\
&= \frac{1 - (1 + \xi \frac{y+u-u_0}{\sigma})^{-\frac{1}{\xi}} - 1 + (1 + \xi \frac{u-u_0}{\sigma})^{-\frac{1}{\xi}}}{(1 + \xi \frac{u-u_0}{\sigma})^{-\frac{1}{\xi}}} \\
&= 1 - (1 + \frac{\xi y}{\sigma + \xi(u - u_0)})^{-\frac{1}{\xi}} \\
&\sim H_{\xi, \sigma + \xi(u - u_0)}(y)
\end{aligned}$$

Les moments d'ordre  $r$  d'une variable  $X$  de loi GPD de paramètre  $\xi$  et  $\beta$  existe pour tout  $\xi < r^{-1}$ . En particulier, l'espérance est donnée par

$$\mathbb{E}[X] = \frac{\sigma}{1 - \xi}$$

donc pour la variable  $X - u \mid X > u$

$$\forall u > u_0 \quad e(u) = \mathbb{E}[X - u \mid X > u] = \frac{\sigma}{1 - \xi} + \frac{\xi(u - u_0)}{1 - \xi}$$

La fonction des excès  $e(u)$  est donc **linéaire** en  $u$  au-delà du seuil  $u_0$ .

En **pratique**, on **choisit** le seuil  $u_0$  dès que la fonction des excès, fonction de  $u$ , est peu près une **droite croissante** à partir de  $u_0$ .



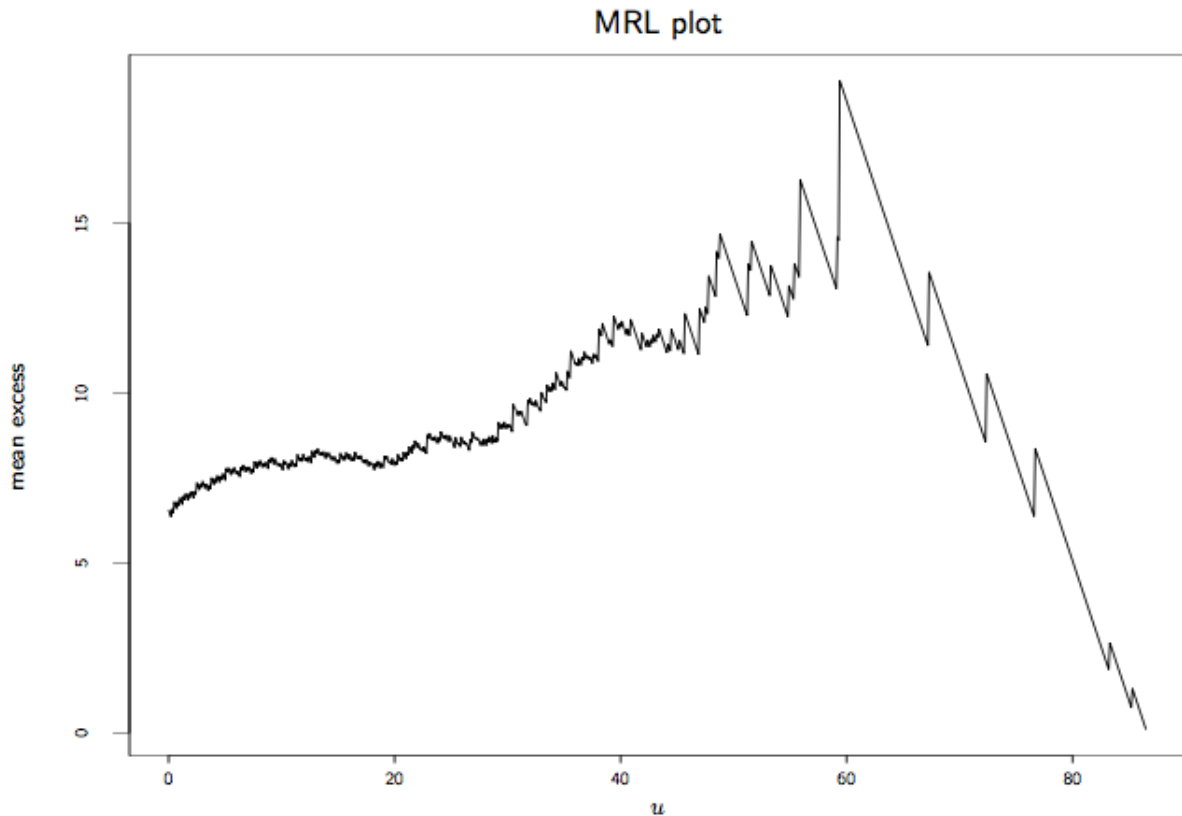


FIGURE E.1 – Exemple de graphique d’espérance excédentaire

Le choix est subjectif néanmoins le graphique nous suggère de choisir le seuil  $u_0 = 30$ . On observe à partir de ce seuil un droite croissante. Il ne faut pas considérer la partie droite du graphique, la variabilité est due à un nombre insuffisant de donnée pour ces seuils très élevés.

## E.2 Graphique de stabilité (Hill, Pickands, ...)

Soit  $(x_1, \dots, x_n)$  l’observation du n-échantillon  $(X_1, \dots, X_n)$ .

Commençons à introduire la définition du domaine d’attraction d’une distribution de la famille de Fréchet.

Une fonction<sup>2</sup>  $l$  **positive** est à **variation lente**. en l’infini si

$$\forall t > 0 \quad \lim_{x \rightarrow +\infty} \frac{l(tx)}{l(x)} = 1$$

---

2. Les fonctions constante, logarithme sont des fonctions à variation lente.

**Proposition E.2.1**

$F_X$  **appartient** au domaine d'attraction d'une distribution de **Fréchet** de paramètre de forme  $\xi$  si et seulement si  $\forall x > 0 \quad 1 - F_X(x) = x^{-\frac{1}{\xi}} l(x)$  où  $l$  est une fonction à **variation lente**.

Il s'agit ici d'estimer le paramètre de forme  $\xi$ .

En utilisant la propriété (E.2.1) on a :

$$\forall t > 0 \quad \lim_{u \rightarrow +\infty} \frac{1 - F_X(tu)}{1 - F_X(u)} = t^{-\frac{1}{\xi}} \quad \lim_{u \rightarrow +\infty} \frac{l(tu)}{l(u)} = t^{-\frac{1}{\xi}}$$

Ceci est équivalent à<sup>3</sup>

$$\lim_{u \rightarrow +\infty} \bar{F}_X(tu) = \bar{F}_X(u) t^{-\frac{1}{\xi}}$$

En posant  $x = tu$  (donc  $x > u$ ) et  $\alpha = \bar{F}_X(u)$  on a

$$\bar{F}_X(x) \underset{+\infty}{\sim} \alpha \left( \frac{x}{\bar{F}_X(\alpha)^{-1}} \right)^{-\frac{1}{\xi}}$$

puis en posant  $p = \bar{F}_X(x)$  on a

$$p^{-\xi} \underset{0}{\sim} \alpha^{-\xi} \frac{\bar{F}_X(p)^{-1}}{\bar{F}_X(\alpha)^{-1}} \Leftrightarrow \bar{F}_X(p)^{-1} \underset{0}{\sim} \bar{F}_X(\alpha)^{-1} \left( \frac{p}{\alpha} \right)^{-\xi}$$

En prenant le logarithme on a  $\ln \bar{F}_X(p)^{-1} \underset{0}{\sim} \ln \bar{F}_X(\alpha)^{-1} - \xi \ln \frac{p}{\alpha}$ .

Si l'on choisit  $\alpha = \frac{k}{n}$  on peut prendre des  $p \in \{\frac{1}{n}, \dots, \frac{k-1}{n}\}$  pour respecter<sup>4</sup>  $p < \alpha$  :

$$\forall i \in \{1, \dots, k-1\} \quad \ln \bar{F}_X\left(\frac{i}{n}\right)^{-1} - \ln \bar{F}_X\left(\frac{k}{n}\right)^{-1} \underset{0}{\sim} \xi \ln \frac{k}{i}$$

Ainsi en prenant la distribution empirique de la fonction de survie  $\bar{F}_X$  on a

$$\forall i \in \{1, \dots, k-1\} \quad \ln \bar{F}_{n,X}\left(\frac{i}{n}\right)^{-1} - \ln \bar{F}_{n,X}\left(\frac{k}{n}\right)^{-1} \underset{0}{\sim} \xi \ln \frac{k}{i}$$

or  $F_{n,X}(x_{(i:n)}) = \frac{i}{n}$  alors on a

$$\forall k \in \{2, \dots, n\} \quad \forall i \in \{1, \dots, k-1\} \quad \ln x_{(n-i:n)} - \ln x_{(n-k:n)} \underset{0}{\sim} \xi \ln \frac{k}{i} \tag{E.1}$$

donc en sommant sur les  $i$  de part et d'autre de  $\underset{0}{\sim}$  on obtient l'estimateur du paramètre  $\xi$

$$\begin{aligned} \hat{\xi}_{k,n} &:= \frac{\sum_{i=1}^{k-1} \ln x_{(n-i:n)} - \ln x_{(n-k:n)}}{\sum_{i=1}^{k-1} \ln \frac{k}{i}} \\ &= \frac{\sum_{i=1}^{k-1} \ln x_{(n-i:n)} - \ln x_{(n-k:n)}}{\ln \frac{k^{k-1}}{(k-1)!}} \end{aligned}$$

3. On note  $\bar{F}_X(\cdot)$  la fonction de survie.

4. Puisque  $x > u$  on a un niveau de confiance  $1 - p$  tel que  $p < \alpha$ .

or si  $k$  est *assez grand* on peut utiliser la formule de Stirling  $k! \underset{+\infty}{\sim} \sqrt{2\pi k} k^k e^{-k}$  et approximer le dénominateur par  $k$  on obtient ainsi l'**estimateur de Hill**

$$\boxed{\text{pour } k \text{ assez grand} \quad \widehat{\xi}_{k,n}^{\text{Hill}} := \frac{1}{k-1} \sum_{i=1}^{k-1} \ln x_{(n-i:n)} - \ln x_{(n-k:n)}}$$

L'estimateur de Hill s'interprète donc comme la *moyenne* des **excès** des  $k - 1$  plus grandes **log**-observations.

L'équation (E.1) permet une détermination **graphique** de l'estimateur  $\xi$ . En effet, pour  $k$  **fixé**, les points

$$\left( \ln X_{(n-i:n)} - \ln X_{(n-k:n)}, \ln \frac{k}{i} \right) \quad \forall i \in \{1, \dots, k-1\}$$

sont **alignés** sur la droite de **pente**  $\xi$ .

En pratique, pour déterminer le seuil  $u_0$ , on trace le graphique de Hill <sup>5\*</sup> représentant l'estimateur de Hill en fonction de  $k$ . On choisit le seuil lorsque l'estimateur est à peu près **stable**.

---

5. Hill-plot

6. \* Exemple du blog d'A. Charpentier [77]

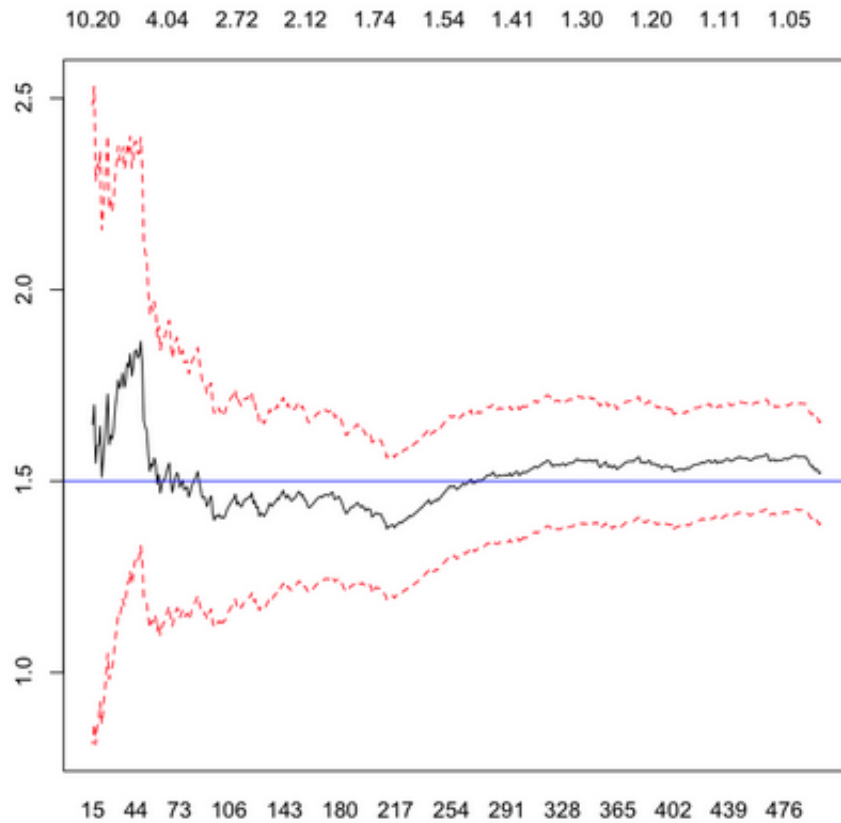


FIGURE E.2 – Exemple de graphique de Hill <sup>6</sup>

Le graphique de Hill nous suggère de choisir un  $u_0$  entre 3 et 4.

Néanmoins l'estimateur de Hill n'est valable que pour  $\xi > 0$ . Le graphique de Pickands permet de prendre en compte les autres cas il est basé sur l'estimateur de Pickands :

$$\text{pour } k \text{ assez grand} \quad \hat{\xi}_{k,n}^{\text{Pickands}} := \frac{1}{\ln 2} \ln \frac{x_{(n-k:n)} - x_{(n-2k:n)}}{x_{(n-2k:n)} - x_{(n-4k:n)}}$$

### E.3 Estimation non-paramétrique du paramètre $\xi$

On utilise par exemple l'estimateur **non-paramétrique** de Hill.

### E.4 Estimation paramétrique du paramètre de dispersion $\sigma$

On utilise un estimateur **paramétrique** construit via le maximum de vraisemblance.

Soient  $(y_{(n-k+1:n)}, \dots, y_{(n:n)})$  l'échantillon d'observation des variables des excès

$$\forall i \in \{1, \dots, k-1\} \quad Y_{(n-i:n)} := X_{(n-i:n)} - X_{(n-k:n)}$$

La vraisemblance de ce  $(k-1)$ -uplet conditionnellement à l'indice de queue  $\xi$ , pour tout seuil  $u$  élevé, est donnée<sup>7</sup> par

$$\begin{aligned} L(\sigma(u); y_{(n-k+1:n)}, \dots, y_{(n:n)} / \xi_{k,n} > 0) &:= \prod_{i=1}^{k-1} H'_{\xi_{k,n}, \sigma(u)}(y_{(n-i+1:n)}) \\ &= \prod_{i=1}^{k-1} \sigma^{-1} \left(1 + \frac{\xi_{k,n}}{\sigma} y_{(n-i+1:n)}\right)^{-\frac{1}{\xi_{k,n}} - 1} \end{aligned}$$

On déduit<sup>8</sup> la log-vraisemblance

$$l(\beta(u); y_{(n-k+1:n)}, \dots, y_{(n:n)} / \xi_{k,n} > 0) = -n \ln \beta - \left(1 + \frac{1}{\xi_{k,n}}\right) \sum_{i=1}^{k-1} \ln \left(1 + \frac{\xi_{k,n}}{\sigma} y_{(n-i+1:n)}\right)$$

L'estimateur de maximum de vraisemblance  $\sigma(u)$  conditionnellement à  $\xi > 0$  est donc donné par

$$\hat{\sigma}(u) := \arg \max_{\sigma(u)} l(\sigma(u); y_{(n-k+1:n)}, \dots, y_{(n:n)} / \xi_{k,n} > 0)$$

On obtient ainsi l'estimateur en dérivant la log-vraisemblance :

$$\frac{\partial}{\partial \sigma} l(\sigma; \cdot) = 0 \Leftrightarrow \hat{\sigma}^{\text{MV}}(u) = \frac{\hat{\xi}_{k,n}}{k-1} \left(1 + \frac{1}{\hat{\xi}_{k,n}}\right) \sum_{i=1}^{k-1} \frac{y_{(n-i+1:n)}}{1 + \frac{\hat{\xi}_{k,n}}{\hat{\sigma}(u)} y_{(n-i+1:n)}}$$

---

7. Par indépendance des variables  $Y_{(n-i:n)}$

8. Par définition de la loi GPD définie en (3.3) page 77.

## Annexe F

# Méthodes numériques d'optimisation

### F.1 Algorithme à évolution différentielle

Cet algorithme a été introduit par Storn et Price en 1997<sup>1</sup> pour résoudre les problèmes d'**optimisation** dont la fonction **objective**  $f$  peut

- présenter des caractéristiques de **non**-linéarité, de **non**-convexité et/ou de **non**-régularité
- posséder **plusieurs** minimum locaux

Il s'agit donc de résoudre le problème<sup>2</sup>

$$x^* := \arg \min \{f(x)\} \quad \forall f : \Omega \subseteq \mathbb{R}^p \mapsto \mathbb{R}$$

L'algorithme est **itératif** :

1. On note pour chaque **génération**  $G$  les vecteurs de paramètres :

$$x_{i,G} := [x_{1,i,G}, \dots, x_{p,i,G}] \quad \forall i = 1, \dots, n$$

Nous avons un  $n$ -échantillon de vecteurs paramètres c'est-à-dire une matrice de taille  $n \times p$ .

2. **Initialisation** :

On définit des bornes sur les paramètres :

$$x_j^L \leq x_{j,i,G=1} \leq x_j^U$$

3. **Mutation génétique** :

---

1. Cf. l'article [37].  
2. Dans le cas présent,  $p$  représente le nombre de paramètres.

Pour le vecteur  $i$  de  $p$  paramètres  $x_{i,G}$  on choisit **aléatoirement** trois vecteurs de paramètres  $(x_{i_1,G}, x_{i_2,G}, x_{i_3,G})$  parmi le  $n$ -échantillon. On note un **vecteur donneur**

$$\tilde{x}_{i,G+1} := x_{i_1,G} + F(x_{i_2,G} - x_{i_3,G})$$

avec  $F$  le **facteur de mutation** une constante choisie dans de l'intervalle  $[0, 2]$ .

#### 4. Recombinaison génétique :

Nous allons considérer le **vecteur d'essai** défini par

$$\tilde{x}_{j,i,G+1} := \begin{cases} \tilde{x}_{j,i,G+1} & \text{si } u_{j,i} \leq \text{CR ou si } u_{j,i} = j \\ x_{j,i,G} & \text{si } u_{j,i} > \text{CR ou si } u_{j,i} \neq j \\ i = 1, \dots, n; j = 1, \dots, p \end{cases}$$

où  $u_{j,i}$  est une réalisation de la loi uniforme  $\mathcal{U}[1, p]$ .

#### 5. Sélection génétique :

$$x_{i,G+1} := \begin{cases} \tilde{x}_{i,G+1} & \text{si } f(\tilde{x}_{i,G+1}) \leq f(x_{i,G+1}) \\ x_{i,G} & \text{sinon} \\ i = 1, \dots, n \end{cases}$$

#### 6. Critère d'arrêt :

Les mutations, les recombinaisons et les sélections génétiques sont effectuées jusqu'à qu'un **critère d'arrêt** soit vérifié :

— le **nombre de génération maximum**  $G_{\max}$  est atteint.

— la fonction objective est *nulle* :

$$|f(x_{i,G+k})| < \epsilon$$

## F.2 Algorithme de Broyden-Fletcher-Goldfarb-Shanno

On appelle **descente** tout vecteur  $d$  de  $\mathbb{R}^p$  tel que le produit scalaire  $\langle \nabla f(x), d \rangle$  soit **négatif**. Par définition de la dérivée, si  $d$  est une direction de descente alors pour tout  $\alpha$  suffisamment petit on a :

$$f(x + \alpha d) < f(x)$$

Cela signifie que **localement** la fonction  $f$  diminue en effectuant un déplacement dans la direction  $d$  avec un pas  $\alpha$ .

Les méthodes de descentes suivent ce **principe** en construisant une suite  $\{x_k\}_{k+1}$  approchant la solution  $x^*$  du problème d'optimisation :

$$x_{k+1} = x_k + \alpha_k d_k.$$

Les méthodes de descentes sont donc déterminées par les paramètres  $d_k$  et  $\alpha_k$ . La **méthode de Newton** consiste à prendre pour direction de descente la direction :

$$d_k := -H_k^{-1} \nabla f_k$$

où  $H_k := \nabla^2 f_k$  est la **hessienne** de la fonction  $f$ .

Mais souvent dans la pratique  $H_k^{-1}$  est très difficile à évaluer, on calcule une approximation du hessien. Broyeden et al. propose l'**approximation**

$$H_{k+1} = H_k + \frac{y_k y_k^\top}{y_k^\top s_k} - \frac{H_k s_k s_k^\top H_k}{s_k^\top H_k s_k}$$

où

$$s_k := x_{k+1} - x_k$$

$$y_k := \nabla f_{k+1} - \nabla f_k$$



# Bibliographie

## - ARTICLES ET PRÉSENTATIONS

- [1] ACPR [2015] : ANALYSE DE L'EXERCICE 2014 DE PRÉPARATION À SOLVABILITÉ II, Analyses et Synthèses, numéro 41, février 2015
- [2] ACPR [2015] NOTICE TECHNIQUE DÉRIVANT LES SCÉNARIOS D'ENVIRONNEMENT DE TAUX BAS DANS LE CADRE DE LA REMISE PRÉPARATOIRE 2015 DE L'ORSA, 25 mai 2015
- [3] K. C. AHLGRIM et al. [2005] : MODELING FINANCIAL SCENARIOS : A FRAMEWORK FOR THE ACTUARIAL PROFESSION.
- [4] A. ALFONSI [2005] : ON THE DISCRÉTISATION SCHEMES FOR THE CIR (AND BESSEL SQUARED) PROCESSES.
- [5] J. ANNAERT et al. [2013] : ESTIMATING THE YIELD CURVE USING THE NELSON-SIEGEL MODEL, A RIDGE REGRESSION APPROACH - International Review of Economics and Finance, Vol. 27, p. 482-496.
- [6] D. BAUER et al. [2008] : ON THE RISK-NEUTRAL VALUATION OF LIFE INSURANCE CONTRACTS WITH NUMERICAL METHODS IN VIEW.
- [7] D. BAUER et al. [2009] : SOLVENCY II AND NESTED SIMULATIONS - A LEAST-SQUARES MONTE CARLO APPROACH.
- [8] P. BERTAIL [2014] : EXTREME VALUES STATISTICS FOR MARKOV CHAINS WITH APPLICATIONS TO FINANCE AND INSURANCE - Extreme value theory and applications in finance, Wiley, Handbook Series in Financial Engineering and Econometrics.
- [9] T. BOLLERSLEV [1986] : GENERALISED AUTOREGRESSIVE CONDITIONAL HETEROSKEDASTICITY - Journal of Econometrics North-Holland, vol. 31, pp. 307-327.
- [10] J. CHRISTENSEN et al. [2010] : THE AFFINE ARBITRAGE-FREE CLASS OF NELSON-SIEGEL TERM STRUCTURE MODELS - Working papers series 2007-20, Federal reserve bank of San Fransisco.
- [11] G. IDABOUK [2008] : MARCHÉS FINANCIERS, MOUVEMENT BROWNIEN, ARBITRAGE ET MARTINGALES : NAISSANCE D'UN CORPUS DE FINANCE MATHÉMATIQUE À PARTIR DE 1973
- [12] F. BONNIN et al. [2011] : APPLICATION DE TECHNIQUES STOCHASTIQUES POUR L'ANALYSE PROSPECTIVE DE L'IMPACT COMPTABLE DU RISQUE DE TAUX
- [13] M-C BRASSIER et al. [2010] : COPULES ET DÉPENDANCE ENTRE LES RISQUES - Séminaire ALTIA
- [14] D. CADOUX et al. [2004] : COPULES ET DÉPENDANCES : APPLICATION PRATIQUE À LA DÉTERMINATION DU BESOINS EN FONDS PROPRES D'UN ASSUREUR NON VIE

- [15] L. DEVINEAU & S. LOISEL [1996] : CONSTRUCTION D'UN ALGORITHME D'ACCÉLÉRATION DE LA MÉTHODE DES SIMULATIONS DANS LES SIMULATIONS POUR LE CALCUL DU CAPITAL ÉCONOMIQUE SOLVABILITÉ II
- [16] L. DEVINEAU & M. CHAUVIGNY [2010] : REPLICATING PORTFOLIOS : TECHNIQUES DE CALIBRAGE POUR LE CALCUL DU CAPITAL ÉCONOMIQUES SOLVABILITÉ II
- [17] F. DIEBOLD & C. LI [2006] : FORECASTING THE TERM STRUCTURE OF GOVERNMENT BOND YIELDS - Journal of Econometrics 130, p. 330-364
- [18] R. ENGLE [1982] : AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTICITY WITH ESTIMATES OF THE VARIANCE OF UNITED KINGDOM INFLATION - Econometrika, Vol. 50, No 4 p. 987-1008
- [19] C. GENEST & R. MACKAY [1986] : COPULES ARCHIMÉDIENNES ET FAMILLES DE LOIS BIDIMENSIONNELLES DONT LES MARGES SONT DONNÉES, The Canadian Journal of Statistics, Vol. 14, N°2, pp. 145-159
- [20] T. J. GEOGHEGAN et al. [1992] REPORT ON THE WILKIE STOCHASTIC INVESTMENT MODEL - Journal of the Institute of Actuaries, Vol. 119, n°2, p. 173-228
- [21] J. HUA [2011] : TERM-STRUCTURE MODELING AND FORECASTING - Risk Professional
- [22] P. KATSIAMPA [1998] : NEW APPROACH TO MODELLING NONLINEAR TIME SERIES : INTRODUCING THE EXPAR-ARCH AND EXPAR-GARCH MODELS AND APPLICATIONS
- [23] R. KAUFMANN et al. [2001] : INTRODUCTION TO DYNAMICAL FINANCIAL ANALYSIS - Astin Bulletin, Vol. 31, n°1, p.213-249
- [24] A. KITTS [1990] : COMMENTS ON A MODEL OF RETAIL PRICE INFLATION - Journal of the Institute of Actuaries, Vol. 117, n°2, p. 407-413
- [25] A. LONGSTAFF & E. SCHWARTZ [2001] : VALUING AMERICAN OPTIONS BY SIMULATION : A SIMPLE LEAST-SQUARES APPROACH
- [26] A. McNEIL & R. FREY [1999] : ESTIMATION OF TAIL-RELATED RISK MEASURES FOR HETEROSCEDASTIC FINANCIAL TIMES SERIES : AN EXTREME VALUE APPROACH
- [27] C. MEYER [2009] : THE BIVARIATE NORMALE COPULA
- [28] C. NELSON & A. SIEGEL [[1987] : PARCIMONIOUS MODELING OF YIELD CURVES - The Journal of Business, Volume 60, Issue N°4, pp. 473-489
- [29] F. PLANCHET & P-E. THEROND [2006] : SIMULATION DE TRAJECTOIRES DE PROCESSUS CONTINUS
- [30] F. PLANCHET [2009] : MODÉLISER L'ACTIF D'UN ORGANISME ASSUREUR : QUELQUES RÉFLEXIONS, Journées d'étude IA/SACEI 15 mai 2009
- [31] F. PLANCHET [2015] : CONSTRUIRE UN GÉNÉRATEUR DE SCÉNARIOS ÉCONOMIQUES EN ASSURANCE, INTRODUCTION - V. 2.9, Prim'Act
- [32] F. PLANCHET & P-E. THEROND [2005] : L'IMPACT DE LA PRISE EN COMPTE DES SAUTS BOURSIERS DANS LES PROBLÉMATIQUES D'ASSURANCE -
- [33] M. ROCCO [2011] : EXTREME VALUE THEORY FOR FINANCE : A SURVEY - Questioni di Economia e Finanza Number 99, Banca d'Italia
- [34] T. RONCALLI, A. NIKEGHBALI & al. [2000] : COPULAS FOR FINANCE A READING GUIDE AND SOME APPLICATIONS - Groupe de Recherche Opérationnelle
- [35] B. SALANIÉ [1999] : GUIDE PRATIQUE DES SÉRIES NON-STATIONNAIRES, Économie et Prévisions, Vol. 137, N° 1, pp. 119-141

- [36] B. SCHWEIZER & E. WOLFF [1981] : ON NON-PARAMETRIC MEASURES OF DEPENDENCE FOR RANDOM VARIABLES. - Ann. Statist. pp. 879-885.
- [37] R. STORN & K. PRICE [1997] : DIFFERENTIAL EVOLUTION - A SIMPLE AND EFFICIENT HEURISTIC FOR GLOBAL OPTIMIZATION OVER CONTINUOUS SPACES - Journal of Global Optimization, pp. 341-359.
- [38] M. VAVRA [2012] : A NOTE ON THE FINITE SAMPLE PROPERTIES OF THE CLS METHOD OF TAR MODELS - Birkbeck Working Papers in Economics and Finance
- [39] A.D. WILKIE [1995] : MORE ON A STOCHASTIC ASSET MODEL FOR ACTUARIAL USE - British Actuarial Journal, Vol. 1, No. 5, pp. 777-964

- **LIVRES**

- [40] G. BOX & G. JENKINS [1976] : TIME SERIES ANALYSIS, FORECASTING AND CONTROL - Prentice Hall.
- [41] BROCKWELL & DAVIS [1991] : TIMES SERIES : THEORY AND METHODS - Springer Series in Statistics.
- [42] R. DAVIDSON et al. [1993] : ESTIMATION AND INFERENCE IN ECONOMETRICS - Oxford University Press
- [43] M. DENUIT & A. CHARPENTIER [2005] : MATHÉMATIQUES DE L'ASSURANCE NON-VIE, TARIFICATION ET PROVISIONNEMENT TOME II - Economica
- [44] DAVISON & HINKLEY : BOOTSTRAP METHOD AND THEIR APPLICATION
- [45] B. EFRON & R. TIBSHIRANI [1993] : AN INTRODUCTION TO THE BOOTSTRAP
- [46] B. EFRON : LE BOOTSTRAP ET SES APPLICATIONS, DISCRIMINATION & RÉGRESSIONS - Traduction de E. Jolivet et R. Hordan
- [47] P. EMBRECHTS et al. [1997] : MODELLING EXTREMAL EVENTS FOR INSURANCE AND FINANCE - Springer Applications of mathematics, Vol. 33
- [48] A. FALEH et al. [2012] : SCÉNARIOS ÉCONOMIQUES ET TECHNIQUES D'ALLOCATION D'ACTIFS, APPLICATIONS AUX ASSURANCES ET AUX FONDS DE PENSION - Economica
- [49] O. KALLENBERG [2002] : FOUNDATIONS OF MODERN PROBABILITY - Springer, Probability and its applications, Second Edition
- [50] S. KOTZ et al. [2000] : EXTREME VALUE DISTRIBUTIONS, THEORY AND APPLICATIONS - Imperial College Press
- [51] R. NELSEN [1999] : AN INTRODUCTION TO COPULAS - Springer, Lecture Notes in Statistics, Vol. 139
- [52] F. PLANCHET et al. [2009] : SCÉNARIOS ÉCONOMIQUES EN ASSURANCE, MODÉLISATION ET SIMULATION - Economica
- [53] L. MARTELLINI & P. PRIAULET [2000] : PRODUITS DE TAUX D'INTÉRÊTS, MÉTHODES DYNAMIQUES D'ÉVALUATION ET DE COUVERTURE - Economica
- [54] Q. GUIBERT et al. [2014] : SOLVABILITÉ PROSPECTIVE EN ASSURANCE MÉTHODES QUANTITATIVES POUR L'ORSA - Economica
- [55] T. RONCALLI [2009] : LA GESTION DES RISQUES FINANCIERS - Economica 2<sup>ème</sup> édition

## - COURS

- [56] P. BERTAIL [2004-a] : BOOTSTRAP IN THE I.I.D CASE - Chapitre 1, Document pour le cours de troisième année.
- [57] P. BERTAIL [2004-b] : SEMIPARAMETRIC BOOTSTRAP FOR TIME SERIES - Chapitre 2 - Document pour le cours de troisième année.
- [58] P. BERTAIL [2007] : MÉTHODES DE BOOTSTRAP ET APPLICATIONS ACTUARIELLES - Formation Caritat.
- [59] A. CHARPENTIER [2014] : COURS DE SÉRIES TEMPORELLES THÉORIE ET APPLICATIONS - Vol 1 et Vol 2 - ENSAE & Université Paris Dauphine
- [60] J-P. FELIX [2010] : "COURS DE MATHÉMATIQUES FINANCIÈRES - MÉTHODE RISQUE-NEUTRE", Université de Picardie Jules Verne - BNP PARIBAS Assurance
- [61] A. MORINEAU [2011] : TUTORIEL BOOTSTRAP ET PROCÉDURE DE RÉ-ÉCHANTILLONNAGE
- [62] R. RAKATOMALALA [2011] : TESTS DE NORMALITÉ, TECHNIQUES EMPIRIQUES ET TESTS STATISTIQUES - V. 2.0, Université Lumière Lyon 2

## - MÉMOIRES D'ACTUARIAT

- [63] K. ARMEL [2010] : STRUCTURE DE DÉPENDANCE DES GÉNÉRATEURS DE SCÉNARIOS ÉCONOMIQUES - EURIA
- [64] C. DENISE-BAILLON [2013] : MODÉLISATION STATISTIQUE DU RISQUE DE TEMPÊTES TROPICALES DANS LE BASSIN ATLANTIQUE NORD - GENERALI – ISFA
- [65] T. GAUTIER [2011] : DÉVELOPPEMENT D'UN GÉNÉRATEUR DE SCÉNARIOS ÉCONOMIQUES AU SEIN D'UNE COMPAGNIE D'ASSURANCE-VIE - ISFA
- [66] M-P. GIROUX [2011] : MODÉLISATION D'UN GÉNÉRATEUR DE SCÉNARIOS ÉCONOMIQUES - PWC – ENSAE ParisTech
- [67] V. LAGENE BRE [2009] : PRISE EN COMPTE DES DÉPENDANCES ENTRE RISQUES PAR LA THÉORIE DES COPULES : IMPACT SUR UN BESOIN EN FONDS PROPRES DE L'ASSUREUR - Winter & associés - EURIA
- [68] G. RAMBARUTH [2003] : A COMPARISON OF WILKIE-TYPE STOCHASTIC INVESTMENT MODELS - Dissertation, Master of Science in Actuarial Management, The City University London

## - TEXTES RÉGLEMENTAIRES

- [69] 2009/138/CE [2009] : DIRECTIVE DU PARLEMENT EUROPÉEN ET DU CONSEIL SUR L'ACCÈS AUX ACTIVITÉS DE L'ASSURANCE ET DE LA RÉASSURANCE ET LEUR EXERCICE (SOLVABILITÉ II) - Journal Officiel de l'Union Européenne
- [70] CEIOPS-DOC-65/10 [2010] : CEIOPS'ADVICE FOR LEVEL 2 IMPLEMENTING MEASURES ON SOLVENCY II : ARTICLE 111 AND 304 EQUITY RISK SUB-MODULE (FORMER CONSULTATION PAPER 69)

- SITES INTERNET

- [71] [http://ec.europa.eu/finance/insurance/solvency/index\\_fr.htm](http://ec.europa.eu/finance/insurance/solvency/index_fr.htm) SITE EN FRANÇAIS DE LA COMMISSION EUROPÉENNE SUR SOLVABILITÉ II Y COMPRIS OMNIBUS II
- [72] <http://acpr.banque-france.fr/solvabilite2.html> : ESPACE DÉDIÉ À LA PRÉPARATION À SOLVABILITÉ
- [73] <http://www.cfoforum.nl/> : EUROPEAN INSURANCE CFO FORUM
- [74] <http://eiopa.europa.eu> : EUROPEAN INSURANCE AND OCCUPATIONAL PENSIONS AUTHORITY
- [75] <http://www.ressources-actuarielles.net> : CE SITE FOURNIT DES RESSOURCES (SUPPORTS DE COURS, ARTICLES, PUBLICATIONS, ETC.) EN ACTUARIAT
- [76] <http://actudactuaires.typepad.com/> : ACTU D'ACTUAIRES - L'ACTUALITÉ DES TECHNIQUES ACTUARIELLES
- [77] <http://freakonometrics.hypotheses.org> : BLOG D'ARTHUR CHARPENTIER - STATISTIQUES ET ACTUARIAT
- [78] <https://cran.r-project.org> THE COMPREHENSIVE R ARCHIVE NETWORK.